

التقويم

الافسسي



صلاح فؤاد سليم



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

﴿ وَقُلْ اَعْمَلُوا فَسَيَرَى اللَّهُ عَمَلَكُمْ وَرَسُولُهُ وَالْمُؤْمِنُونَ ﴾

صدق الله العظيم

التقويم النفسي

التقويم النفسي

إعداد

صلاح فؤاد سليم

الطبعة الأولى

2006 م - 1425 هـ



مكتبة المجتمع العربي للنشر والتوزيع

سليم، فؤاد

التقويم النفسي / إعداد فؤاد صلاح سليم. - عمان: مكتبة
المجتمع العربي، 2005.

() ص.

ر.إ: (2005/4/708).

الواصفات: /الختبارات الشخصية// الإرشاد النفسي// علم النفس
الفردي// الشخصية// علم النفس/

تم إعداد بهالفت الفهرسة والتصنيف الأولية من قبل دائرة المكتبة الوطنية

حقوق الطبع محفوظة للناسخ

Copyright ©
All Rights reserved

الطبعة الأولى

2006 م - 1426 هـ



مكتبة المجتمع العربي للنشر والتوزيع

عمان - شارع الملك حسين - مجمع الفحيص التجاري

تلفاكس 4632739

ص.ب. 8244 عمان 11121 الأردن

المحتويات

الوحدة الأولى

9 مبادئ القياس النفسي
---	---------------------------

الوحدة الثانية

67 اختبارات الذكاء العام
----	-----------------------------

الوحدة الثالثة

111 اختبارات القدرة المتخصصة
-----	--------------------------------

الوحدة الرابعة

141 قياس الشخصية عن طريق الأحكام
-----	------------------------------------

الوحدة الخامسة

157 قياس للشخصية (وسائل التقرير الذاتي)
-----	---

الوحدة السادسة

183 اختبارات التحصيل
-----	------------------------

215 المراجع
-----	---------------

مبادئ القياس النفسي

مقدمة:

لواقع أن المعلومات التي تتوافر لنا عن المقياس ذاته: هل هو من نوع مقياس للنسبة أو المسافة أو الرتبة، وعن موضوعيته ومدى تحرره من الأحكام الذاتية، وعن خصائصه: هل هو من النوع " الكلي " أو " التخليقي "، وعن طبيعة البيانات التي يزودنا بها: هل هي تقارير الآخرين أو استجابات مباشرة، ومن مدى تقنين شروط إعطائه، لهذه المعلومات جميعاً تعيننا كثيراً في الحكم على أداة التقويم. إلا أن تقويمنا لأداة التقويم لا يكتمل إلا إذا تولفت لدينا مجموعة أخرى من المعلومات عن مدى دقة البيانات التي نحصل عليها بهذه الأداة، ومدى صلاحية الأداة ذاتها، وبعبارة أخرى ثبات الاختبار وصنقه.

ثبات الاختبارات Test reliability:

يقصد بثبات الاختبارات، وأحوال التقويم عادة، أن تكون على درجة عالية من الدقة والإتقان والاتساق والاطراد فيما تزودنا به من بيانات عن سلوك المفحوص.

ويستخدم مفهوم الثبات بمعناه الواسع ليدل على مدى اعتماد الفروق الفردية في درجات الاختبار على أخطاء الصدفة المتضمنة في القياس من ناحية وعلى الاختلافات الحقيقية الأصلية في الصلة أو الخاصية السيكولوجية موضع الاهتمام من ناحية أخرى. ويمكن صياغة المشكلة صياغة فنية على النحو الآتي: ما هي النسبة في التباين الكلي لدرجات الاختبار التي يمكن اعتبارها من

نوع تبليين الخطأ error variance وبالطبع فإن ما نعتبره من نوع تبليين الخطأ في موقف معين يمكن اعتباره تبليناً حقيقياً في موقف آخر. فمثلاً إذا كان اهتمامنا متجهاً نحو دراسة تذبذب المزاج فإن التغيرات اليومية في درجات اختبار يقيس الانشراح والانتفاض تصبح مرتبطة بغرض الاختبار نفسه وتكون بالتالي جزءاً من التبليين الحقيقي في الدرجات. أما إذا كان الاختبار يهدف إلى قياس الخصائص الأكثر استقراراً في الشخصية تصبح هذه التغيرات اليومية بذاتها من نوع تبليين الخطأ.

وعلى ذلك يمكننا أن نعرف تبليين الخطأ بأنه يتعلق بالشروط والظروف التي لا تتصل من قريب أو بعيد بغرض الاختبار. ولذلك فحينما يحول المدرس أو المرشد النفسي أو الأخصائي النفسي أن يحافظ على توحيد الظروف الاختبارية وتقنياتها وذلك بضبط مكان الاختبار وتعليماته وزمنه وعلاقته بالمختبرين وغيرها من العوامل فإنه بذلك يقلل من تبليين الخطأ ويجعل درجة الاختبار أكثر ثباتاً واستقراراً. ومع ذلك فلا يمكن اعتبار الاختبار أداة ثابتة ثباتاً كاملاً يصل إلى درجة ثبات مقاييس الظواهر الطبيعية (كالمتر للأطوال - الكيلو جرام للأوزان). ولذلك لا بد أن يحدد صانعو الاختبارات مقدار ثباتها في دراسة التعليمات بحيث يكون دليلاً للأخصائيين عندما يطبقون هذه الاختبارات في ظروف معيارية مقلدة.

وبالطبع يمكن أن توجد أنواع من ثبات الاختبارات يحدد الظروف التي يمكن أن تؤثر في درجات هذه الاختبارات ولا تكون كلها أو بعضها غير متعلقة بغرض الاختبار. إلا أن المصادر الرئيسية التي يمكن أن نرجع إليها تبليين

الخطأ والتي تحدد للمقاييس العامة لثبات الاختبارات يمكن تصنيفها في ضوء الطرق المختلفة لحساب الثبات.

(أ) إعادة الاختبار

من المصادر الأساسية لتباين الخطأ وما يتمثل في التغيرات العشوائية التي تتألب الأفراد من جلسة اختبار إلى أخرى. وقد تنتج هذه التقلبات عن ظروف تطرأ على الموقف الاختباري ولا يمكن التحكم فيها مثل التغيرات الفجائية في الجو أو حدوث ضوضاء مفاجئة (وغيرها من مشتتات الانتباه). وقد تنتج عن تغيرات في ظروف المفحوص نفسه كما تتمثل في حالات المرض والتعب والاضطراب الانفعالي والقلق وغيرها. وهذه العوامل تدل على المدى الذي تتأثر به درجات الاختبار بالتقلبات اليومية للعشوائية في المفحوص أو في ظروف الموقف الاختباري.

ومن الواضح أن درجة الاستقرار الوتقي Temporal Stability تعتمد إلى حد ما على طول الفترة الزمنية التي تنقضي والتي تقدر بعدها ثبات الاختبار. وتوجد أمثلة عديدة لاختبارات لها ثبات مرتفع إذا قدر هذا للثبات بعد فترات زمنية قصيرة (أيام قلائل أو أسابيع قلائل) ولكن ثباتها ينخفض انخفاضاً شديداً عندما تمتد هذه الفترات إلى عشر سنوات أو خمسة عشر سنة. وكثير من اختبارات نكاه أطفال ما قبل المدرسة لها ثبات مرتفع في فترة ما قبل المدرسة ولكنها تفشل في التنبؤ بنكاه الأطفال في مرحلة الطفولة المتأخرة أو للمراهقة أو للضحج بسبب ما يطرأ على الطفل من تغيرات ارتقائية في النواحي العقلية.

ولا شكل في أن التنبؤات ذات المدى القصير (والتي تحدث خلال فترات تتراوح بين ساعات قلائل وأشهر قلائل) يمكن اعتبارها من نوع تباين الخطأ في درجة الاختبار. وإذ ذلك ففي الممارسة العملية عندما نقوم بتقدير مثل هذا للووع من الثبات لا بد من أن تكون الفترة الزمنية قصيرة، وتكون أقصر عند اختبارنا للأطفال للصغار منها عند اختبارنا للكبار بسبب سرعة معدل النمو في الطفولة عنه في مراحل النمو الأخرى. وبالنسبة لأي نوع من المفحوصين يجب ألا تزيد فترة ما بين الاختبارين عن ستة أشهر بحال من الأحوال.

ومعنى ذلك أن التغيرات التي تحدث في الأداء النسبي للمفحوصين بعد فترات زمنية طويلة يمكن اعتبارها ضرباً من النمو والتقدم في السلوك وليس مجرد تغير عشوائي، بل قد تنتمي هذه للتغيرات إلى ميدان من السلوك أوسع نطاقاً مما يتناوله الأداء الاختباري ذاته. ومن ذلك أن المفحوص قد يتغير أدائه في اختبارات الاستعداد المدرسي أو الفهم الميكانيكي أو الحكم الفعلي بعد فترة زمنية طويلة عشر سنوات بسبب تعرض المفحوص لخبرات غير عادية في هذه الفترة. بعبارة أخرى يمكن للفرد أن يتغير مركزه النسبي في الاختبار بعد هذه الفترة صعوداً أو هبوطاً بالنسبة لأقرانه من بنفس العمر بسبب ظروف خاصة تعود إلى المنزل أو المدرسة أو البيئة الاجتماعية أو بسبب المرض أو الاضطراب الانفعالي.. إلخ.

ويجب أن نشير إلى أن الوظائف النفسية ذاتها قد تختلف فيما بينها في مدى وما تظهره من ذنبات وتقلبات وقتية فمثلاً قد تتأثر وظيفة سلوكية مثل تران الحركات الدقيقة للأصابع بأي تغير ولو كان طفيفاً في ظروف المفحوص بينما نجد أن سلوكاً آخر مثل الاستدلال لا يسهل التأثير فيه بالمؤثرات

المعارضة. ولذلك إذا أردنا أن نحصل على تقدير دقيق لاتزان الأصلاح فإننا نحتاج إلى إعادة الاختبارات مرات عدة في أيام متوالية، بينما تكفي جلسة اختبار واحدة لقياس الاستدلال. ومعنى ذلك أنه من الضروري تحليل أعراض الاختبار وطبيعة السلوك الذي يدعى الاختبار أنه قادر على التنبؤ به تحليلاً دقيقاً وصيقاً قبل اختبار أنسب الطرق لأصحاب ثباته.

ويتعد معامل الثبات الذي يتأثر بتباين الخطأ الناتج عن التقلبات الزمنية بطريقة إعادة الاختبار بصورته ذاتها على نفس الأفراد في مناسبات مختلفة، ويسمى المعامل الناتج (وهو في حقيقته معامل ارتباط بين مجموعتي نتائج الاختبار في كل مناسبتين طبق فيهما) معامل الاستقرار Coefficient of stability. ورغم بساطة هذه الطريقة فإنها تتضمن بعض المشكلات. منها مثلاً أن الممارسة قد تنتج قرراً من التنصيص في درجات إعادة الاختبار. ومنها أيضاً أنه لو كانت الفترة بين موقف الاختبار وإعادة تنصيصه قصيرة فإن المفحوص قد يستدعي بعض استجاباته السابقة، أو بعارة أخرى قد نحصل على نفس النمط من الإجابة من حيث الوضع النسبي للصواب والخطأ ولكن بسبب الوظيفة السلوكية موضع القياس، وبالتالي لا تكون الدرجات التي حصلنا عليها من موقعي الاختبار مستقلة بعضها عن بعض كما أن معامل الارتباط بين مجموعتي النتائج قد يكون مرتفعاً ارتفاعاً زائفاً. وكذلك من مشكلات طريقة إعادة الاختبار أن طبيعة الاختبار ذاتها تتغير بسبب الإعادة وخاصة إذا كانت أسئلة الاختبار تتطلب استخدام الاستدلال أو التحقق. ففي مثل هذه الأسئلة إذا قام المفحوص باكتشاف العلاقة أو معرفة المبدأ المتضمن في المسألة أو توصل إلى حل المشكلة، فإنه يمكن أن يستعيد الإجابة الصحيحة في المستقبل دون المرور بخطوات التفكير الوسيطة. وعلى ذلك فإن هذه الطريقة لا تصلح إلا مع

الاختبارات التي تتأثر تأثيراً كبيراً بالإعادة أو التكرار مثل اختبارات التمييز الحسي والاختبارات الحركية. أما بالنسبة لمعظم الاختبارات للنفسيه فإن طريقة إعادة الاختبار ليست ملائمة لها.

(ب) الصور المتكافئة:

من المألوف لدى الكثير من الطلاب أن يدخلوا الامتحان " فيسعدهم الحظ " بعدد كبير من الأسئلة يتناول الموضوعات التي اهتموا بدراستها ومذاكرتها أكثر من غيرها، وقد يحدث للعكس في كثير من الأحيان. هذا الموقف يدل على مصدر آخر من مصادر تباین الخطأ في درجات الاختبار وهو مدى اعتماد درجة المفحوص في الاختبار على عوامل تتعلق بطريقة اختيار الأسئلة. فإذا افترضنا أن قام باحث آخر بإعداد اختبار تتوافر فيه نفس مواصفات الاختبار الأول فبلى أي حد تختلف درجة المفحوص في الاختبارين ؟ الواقع أن الفرق في الدرجة التي يحصل عليها نفس الأفراد في هذين الاختبارين هو نوع الثبات التي يتأثر بحيلة الأسئلة التي يتكون منها كل من الاختبارين. فمثلاً إذا افترضنا أن باحثاً قام بإعداد اختبار لقياس المعلومات الميكانيكية يتكون من 50 سؤالاً، ثم قام باحث آخر بإعداد اختبار مكافئ " أي يتكون من 50 سؤالاً كذلك لنفس الغرض ". ولنفرض أن أسئلة كل من الاختبارين قد صيغت بنفس الدرجة من المهارة بحيث تتناول نفس المستوى من الصعوبة، فإنه تبعاً لعوامل الصدفة البحتة في الخبرات الماضية للمفحوصين نجد أن للصعوبة النسبية لمجموعتي الأسئلة سوف تختلف من شخص لآخر، وعلى ذلك فإن الاختبار الأول قد يحتوي على عدد كبير من الأسئلة لم يألّفه المفحوص (أ) في

حين تكون لُفته بأسئلة الاختبار التالي أكبر. وقد يحدث العكس بالنسبة للمفحوص (ب).

ويسمى معامل الثبات في هذه الحالة بمعامل التكافؤ أو التساوي Coefficient of equivalence، ويدل على مدى الارتباط بين درجات اختبار، ومقياس آخر متكافئ معه يعطى في نفس الوقت. ويمكن أن تحصل عليه بإعطاء صورتين من الاختبار في وقت متقارب جداً، بشرط أن تكون هاتان الصورتان متعادلتين متكافئتين، أي تقومان نفس الصفات العامة للأدائين بنفس الطريقة. ولكن، كما سبق أن أوضحنا، قد تؤثر للصفات الخاصة تأثيراً مختلفاً على الاختبارين مثل المعلومات الخاصة التي قد تساعد في الإجابة على سؤال معين من أسئلة أحد الاختبارين ولكنها لا تساعد في الاختبار الآخر.

وتباين الخطأ المتضمن هنا يدل على التغيرات في الأداء من مجموعة معينة من الأسئلة إلى مجموعة أخرى، أي الخطأ الذي يرجع إلى عينة الأسئلة. ولكنه لا يتضمن تباين الخطأ الذي يعود إلى العوامل للوقتية، والذي سبق أن تحدثنا عنه، والذي يتمثل في التغيرات في الأداء من موقف اختباري إلى آخر. ويمكن الجمع بين هذين النوعين من تباين الخطأ إذا قمنا بتطبيق الصورتين المتكافئتين من الاختبار في مناسبتين مختلفتين " أي لا نطبقهما في نفس الوقت ". ويمكن بهذه الطريقة تقادي بعض مشكلات طريقة إعادة الاختبار نفسه. ويكون معامل الثبات هو معامل الارتباط بين درجات صورتين الاختبار، وفي هذه الحالة يدل على كل من الاستقرار الزمني واتساق أو إطراد الاستجابة لعينات مختلفة من الأسئلة. ويمكن أن نسمي هذا المعامل معامل الاستقرار والتكافؤ معاً.

وعند إعداد الصور المكافئة من الاختبارات يجب أن نتأكد من أن هذه الصور فعلاً متكافئة. والصور المكافئة هي في جوهرها اختبارات تم إعدادها كل على حدة وبطريقة مستقلة بحيث تتوافر فيها نفس الموصفات، أي أن تحتوي على نفس العدد من الأسئلة، وأن تكون صياغة هذه الأسئلة متماثلة، وأن تتضمن الأسئلة محتوى واحد، وأن تتعامل الأسئلة في مستويات الصعوبة، وأن تتفق الاختبارات في جميع المظاهر الأخرى مثل التعليمات والزمن والأمثلة للتوضيحية والشكل العام. ولا يمكن أن تعزى للفروق في الدرجات من اختبار إلى آخر إلى تباين الخطأ الذي يرجع إلى عينة محتوى الاختبار إلا إذا كانت للصورتان متكافئتين أو متعادلتين حقاً.

ومن المؤكد أن حاجتنا إلى الصور المتكافئة من الاختبارات تتعدى مجرد تحديد ثبات الاختبار. فالصور البديلة تزيد كثيراً في الدراسات التي نجربها حول آثار هامل تجريبي متوسط معين على الأداء الاختباري، كما يفيد استخدام الصور البديلة في إعطاء وسيلة لخفض احتمال للتدريب أو للفشل.

إلا أن أهم صعوبات هذه الطريقة تتمثل في عدم إمكان إعداد صور متكافئة لكثير من الاختبارات لأن ذلك يضاعف الجهد والوقت والتكلفة المبذولة في بناء الاختبار. ومن المعروف أن إعداد اختبار نفسي واحد يتطلب جهداً عظيماً، فما بالك بإعداد أكثر من اختبار لقياس نفس السمة.

(ج) لتجزئة النصفية:

بدل تجانس الأسئلة عن مدى اتساق وإطراد أداء المفحوص على جميع الأسئلة التي يتكون منها الاختبار. فإذا كان لدينا اختباران تتكون أسئلة الأول

منهما من عمليات ضرب فقط بينما تتكون أسئلة الاختبار الثاني من العمليات الحسابية الأساسية الأربعة فإن الاختبار الأول سوف يكون أكثر إتساقاً وإطراداً بين أسئلته من الاختبار للثاني، لأنه في الاختبار الثاني (وهو اختبار غير متجانس الأسئلة) قد يحسن بعض المفحوصين الأداء في عمليات للطرح مثلاً بينما يسيئون أداء باقي العمليات.

لذا زاد حجم تجانس الأسئلة، كأن يتكون الاختبار مثلاً من 40 سؤالاً كل عشرة منها تتناول معاني المفردات والعلاقات المكافئة والاستدلال الحسابي والسرعة الإدراكية على التوالي فقد لا تكون هناك علاقة أو قد تكون العلاقة ضئيلة للغاية بين أداء المفحوص على هذه الأنماط المختلفة من الأسئلة.

ومن ذلك نتبين أن درجة الشخص في اختبار ما تكون أكثر وضوحاً عندما نحصل عليها من اختبارات متجانسة التكوين نسبياً. وتبرز مشكلة وضوح معنى للدرجة الاختبارية إذا عدنا إلى المثال الأخير الذي ذكرناه على الاختبار غير المتجانس. فإذا حصل المفحوصان (أ، ب) على درجة مقدارها 20 في هذا الاختبار فهل يمكننا أن نستنتج من ذلك أن أداءهما متساو تماماً في هذا الاختبار ؟ الإجابة بالنفي دون شك. فقد يلجج المفحوص (أ) مثلاً في الإجابة على أسئلة المفردات والسرعة الإدراكية ويفضل تلمساً في الإجابة على أسئلة الاستدلال الحسابي والعلاقات المكافئة. أما المفحوص (ب) فقد يحصل على درجته من الإجابة على خمسة أسئلة السرعة الإدراكية و خمسة أسئلة للعلاقات المكافئة و عشرة أسئلة للاستدلال الحسابي ويفضل تماماً في الإجابة على أسئلة المفردات. وهناك ما لا حصر له من الاحتمالات في مثل هذا النوع من الاختبارات غير متجانسة التكوين. ومعنى ذلك أن درجة المفحوص في مثل هذا الاختبار لها

معان مختلفة، بينما يكون معناها أكثر وضوحاً في الاختبارات المتجانسة (أو على حد تعبير ثرستون الاختبارات ذات التكوين العملي البسيط).

ورغم أن الدرجة 20 في اختبار متجانس الأسئلة (معاني المفردات مثلاً) قد لا يكون لها معنى واحد تماماً في جميع الحالات، وإنما قد تغطي أن المفحوص لجلب إجابة صحيحة على المفردات العشرين الأولى تقريباً إذا كانت هذه المفردات مرتبة درجة الصعوبة، وربما قد يترك كلمتين أو ثلاث من الكلمات السهلة أو قد يفشل في الإجابة عليها، وقد يلجج في الإجابة على كلمتين أو ثلاث أصعب أو أبعد من الكلمة العشرين في ترتيب الصعوبة. ومع ذلك فإن هذه الاختلافات الفردية قليلة وهشة إذا قررت بذلك التي نلاحظها في الاختبارات الأقل تجانساً.

وهكذا بعد عدم تجانس أسئلة الاختبار مصدراً لتباين الخطأ، وبالتالي يؤثر في ثبات الاختبار. ولكن هل يختلف تجانس الأسئلة كمصدر لتباين الخطأ عن المصدر الذي سبق للحديث عنه وهو ملامحة عينة المحتوى عموماً ؟ الواقع أن الفرق بين هذين للمدين يمكن أن يتضح إذا أعطينا مثلاً متطرفاً. لنفترض أن هناك اختباراً يقيس كل سؤال فيه عملية عقلية أو وظيفة نفسية مختلفة ولا يرتبط أحى ارتباطاً بالأسئلة الأخرى التي يقيس كل منها عمليات أو وظائف أخرى. ولا شك أنه في مثل هذه الحالة يمكن إعداد صورة متكافئة من هذا الاختبار يحتوي على نفس أنواع الأسئلة وتوزيعها على النحو الموجود في الصورة الأولى. وعندما نعطي الصورتين المتكافئتين قد نتفق درجتهما لتتفق كاملاً مما يدل على الوجهة النظرية على معامل ثبات مرتفع لا يتأثر تأثيراً كبيراً بعينة الأسئلة كمصدر لتباين الخطأ. ولكن عندما نهتم بالثبات في ضوء تجانس الأسئلة

نجد أن هذا المعامل يصل إلى الصفر أو قرب إليه، حيث أن إطراد الأداء من سؤال إلى آخر في أي من صورتَي الاختبار قد لا يكون له دلالة أفضل من مستوى الصدفة.

ويسمى معامل الثبات الذي يتأثر بتجانس الأسئلة معامل الاتساق الداخلي، ويمكن للحصول عليه من تطبيق واحد لصورة واحدة من الاختبار، ثم اللجوء إلى ما يسمى طريقة التجزئة للنصفية. وبها يمكن الحصول على درجتين لكل فرد وذلك بتقسيم كل اختبار إلى نصفين يمكن للمقارنة بينهما ثم حساب معامل الارتباط بينهما. ومن الواضح أن حساب الثبات بالتجزئة للنصفية، أو بعبارة أخرى معامل الاتساق الداخلي، لا يعطي مقياساً للتجانس الكلي للمقياس لأنه يقسم أسئلة الاختبار إلى نصفين ولا تتناول هذه الأسئلة سؤالاً موالاً.

وأخطر المشكلات التي تواجهها طريقة التجزئة للنصفية هي كيف نجزي الاختبار للحصول على أفضل قسمين للمقارنة. هناك عدة طرق لقسم الاختبار وخاصة إذا علمنا أن معظم الاختبارات لا يمكن تجزئتها إلى نصفين صالحين للمقارنة بينهما بمجرد تتبع التسلسل الرقمي للأسئلة، وذلك بسبب الفروق في طبيعة الأسئلة ومستوى صعوبتها بالإضافة إلى الآثار التراكمية للحساس في الأداء أو العمل والممارسة والتعب والمال وغيرها من العوامل التي تختلف في بداية الاختبار عنها في نهايته. ولذلك فإن إحدى الطرق الدقيقة للموضوعية للحصول على قسمين متناظرين في الاختبار هي تحديد مستوى صعوبة كل سؤال وذلك بإيجاد النسب المئوية للأفراد الذين أجابوا عليه إجابة صحيحة، ثم توزع الأسئلة على قسمي الاختبار تبعاً لتكافؤ مستوى الصعوبة ولتشابه المضمون. ثم نطبق الاختبار على مجموعة أخرى من المفحوصين غير تلك

التي طبقنا عليها الاختبار لحساب مستويات الصعوبة، وحصل على درجات الأفراد لتقدير الثبات أو معامل الاتساق الداخلي.

ومن الطرق الصالحة لمعظم الأغراض وفي نفس الوقت لا تستغرق وقتاً طويلاً ولا تتطلب جهداً زائداً طريقة تجزئة الاختبار إلى نصفين أحدهما يضم الأسئلة ذات الأرقام الزوجية والآخر يضم الأسئلة ذات الأرقام الفردية، وبخلفة إذا كانت الأسئلة مرتبة في الاختبار تبعاً لمستويات الصعوبة، إلا أننا يجب أن نحذر أنه في حالات مجموعات الأسئلة التي تتعلق بمشكلة واحدة مثل الأسئلة التي تتلو فقرة معينة في اختبار للقراءة فإن هذه الأسئلة جميعها يجب أن توضع في جزء من الاختبار لأننا لو نصفنا مجموعة الأسئلة من هذا النوع على جزئي الاختبار فإن التشابه بين الجزئين سوف يزداد زيادة مصطنعة أو زائفة حيث أن أي خطأ في فهم المشكلة قد يؤثر في أسئلة نصف الاختبار.

وقد أوضحنا أننا نحصل على معامل الاتساق الداخلي بحساب معامل الارتباط بين درجات نصف الاختبار ولكننا يجب أن نشير إلى أن مثل هذا المعامل يعطينا ثبات نصف الاختبار فقط وليس الاختبار كله. وهذه الخاصية لا توجد بالطبع في الطرق السابقة كإعادة الاختبار أو الصور المتكافئة. ونحن نعلم أنه إذا تساوت الظروف الأخرى فإن الاختبار الأطول من حيث عدد أسئلته يكون أعلى ثباتاً، لأننا لو حصلنا على عينة أكبر من السلوك نصل إلى مقاييس أكثر دقة وأكثر استقراراً ويمكن تقدير الأثر الذي تحدثه إطالة الاختبار أو اختصاره على معامل الثبات باستخدام معادلة سبيرمان براون المشهورة وهي:

$$r = \frac{r_{11}}{1 + (n-1)r_{11}}$$

حيث أن ر = معامل الثبات الذي نريد الحصول عليه.

رأ - معامل الثبات الذي حصلنا عليه.

ن - عدد مرات إطلاعة الاختبار أو اختصاره أي أنه إذا زاد عدد الأسئلة من 25 - 100 سؤال فإن (ن) تكون 4 أما إذا اختصر من 60-30 سؤال فإن (ن) تصبح $\frac{1}{2}$. ومعنى ذلك أن هذه المعادلة حينما تطبق على معامل الثبات المقدر بطريقة التجزئة النصفية فإنها تتضمن إطلاعة الاختبار إلى النصف أي تصبح المعادلة على النحو الآتي:

$$r = \frac{r_2}{r_1 + 1}$$

إلا أن معادلة سبيرمان براون تفترض أن للتغيرات في درجات نصفي الاختبار متساوية تماماً وهو افتراض يصعب في كثير من الأحوال تحقيقه حتى ولو بدأ نصفاً الاختبار على درجة كبيرة من التكافؤ. ولذلك فقد يستعاض عن ذلك باستخدام معادلة جتمان (12) التي تستبعد هذا الافتراض ويمكن صياغتها على النحو الآتي:

$$r_{1-2} = \left[\frac{r_1^2 + r_2^2}{r_1^2} - 1 \right]$$

وفي هذه المعادلة نجد أن ع أ و ع ب - الاحرفات المعيارية لتصفي الاختبار .

ع ك - الاحراف المعيارية للدرجات الكلية
في الاختبار .

ومن الواضح أن هذا الإجراء لا يتطلب حساب معامل الارتباط .

(د) تحليل التباين

تصلح طريقة التجزئة النصفية لحساب معامل الاتساق الداخلي (تجانس نصفي الاختبار)، أما بالنسبة لموضوع تجانس الأسئلة ذاتها فإن بعض العلماء يرون أنه لا يدخل في نطاق مناقشة موضوع ثبات الاختبارات وإنما هو من خصائص الاختبارات الجيدة ولا يمكن إدراجه تحت أي من المفاهيم التقليدية للثبات. ومع ذلك فلا يمكن استبعاده تماماً من مناقشة للثبات حيث أنه يتعلق بمدى استقرار استجابات المفحوص. على أسئلة الاختبار واحداً بعد الآخر. وعلى كل فإن مفهوم التجانس الكلي يجب أ، تميزه تمييزاً واضحاً عن غيره من صور الثبات التي سبق لنا مناقشتها. وعادة ما يسمى " معامل الاتساق داخل أسئلة الاختبار " Interitem Consistency وهو مصطلح آخر قد لا يعني نفس ما يعنيه مصطلح معامل الاتساق الداخلي. فمعامل الاتساق داخل الأسئلة يشير إلى التجانس الكلي للاختبار بالمعنى المباشر، أما معامل الاتساق الداخلي فهو أقرب إلى معنى التكافؤ كما يتمثل في الصور المتكافئة للاختبار. ومعنى ذلك أنه قد يوجد لدينا اختباران لهما معاملات ثبات مرتفعة حسبت إما بالصور للمكافئة أو بالتجزئة النصفية ولكنهما قد يختلفان في المعاملات الخاصة بالاتساق داخلي الأسئلة إذا اختلفا في درجة تجانس الأسئلة التي يتكون منها صورتا

الاختبار أو نصفاء. والواقع أن الفرق بين معامل الاتساق الداخلي (الذي يحسب بالتجزئة للنصفية كما سبق أن أوضحنا) وبين معامل الاتساق داخل الأسئلة يمكن اختباره دليلاً على مدى التجانس أو التباين في أسئلة الاختبار.

وأشهر طرق لحساب معامل اتساق الأسئلة طريقة كيودر ريتشاردسون وهذه للطريقة تتشابه مع طريقة حساب معامل الاتساق للدخلي في أنها تستخدم إجراء اختبارياً واحداً للاختبار الواحد، ولكنها بدلاً من أن تتطلب تجزئة الاختبار إلى نصفين متكافئين فإنها تعتمد على فحص أداء الأفراد على كل سؤال على حدة ولذلك فهي تنتمي إلى طرق تحليل التباين بالمعنى العلم، وأشهر المعادلات الإحصائية التي تستخدم في هذا الصدد هي:

$$r = \frac{c^2 - \text{مجموع } x}{c^2} \left(\frac{n}{1 - n} \right)$$

حيث أن r = معامل ثبات الاختبار.

n = عدد الأسئلة التي يتكون منها الاختبار.

c = الانحراف المعياري لدرجات الاختبار.

مجموع x = يحسب بإعداد نسبة الأفراد الذين أصابوا في إجاباتهم على كل سؤال ونسبة الذين أخطئوا ثم نضرب النسبتين في بعضهما ثم نجمع حواصل الضرب بالنسبة لجميع الأسئلة.

ومن الواضح أن معامل الاتساق داخل الأسئلة كما يحسب بطريقة كيودر ريتشاردسون هو في الواقع متوسط معاملات الاتساق للدلطي الناتجة عن عدد من التجزئات للاختبار. أما معامل الاتساق الداخلي فيعتمد على تجزئة تهدف إلى الحصول على مجموعتين متساويتين أو متكافئتين من الأسئلة. وعلى ذلك إذا لم تكن أسئلة الاختبار على درجة كبيرة من التجانس فإن معامل الاتساق داخل الأسئلة يصبح أقل من معامل الاتساق للدلطي.

(هـ) الممتحنون والمصححون:

. بعد أن نستبعد الأنواع الثلاثة السابقة من تبين الخطأ من التباين للكلية (الدرجة التي يحصل عليها المفحوص في الاختبار) فإن التباين للمتبقي يرجع إلى الفروق الفردية الحقيقية لموضوع الاهتمام، وبعضه الآخر إلى عوامل أخرى غير متعلقة ولكن يمكن ضبطها تجريبياً. فمثلاً لا يسجل الباحثون أخطاء للقياس الناتجة عن تطبيق الاختبار في ظروف مشقة للانتباه، أو في زمن أطول أو أقصر مما يحدده كتيب الاختبار، أو بإعطاء تعليمات مختلفة عن التعليمات المحددة لأن هذه الأخطاء جميعاً يمكن استبعادها والتغلب عليها في الموقف الاختباري نفسه. ولذلك فالحسن لا نجد في العادة معاملات ثبات خاصة تشير إلى "تباين الخاص بتشتت الانتباه" أو "تباين اختلاف الزمن" أو "تباين اختلاف التعليمات" ... إلخ. ثم إننا نجد أن معظم الاختبارات تحدد إجراءات مقننة للتطبيق والتصحيح بحيث نفترض مع وجود هذه الإجراءات والالتزام بها أن "ثبات الممتحن" و "ثبات المصحح" يصل إلى درجة كبيرة من الكفاءة تجعل الاختبار مفيداً في الأغراض التطبيقية والعملية وبالتالي لا تكون في حاجة إلى تقدير مثل هذا النوع من الثبات. ويصدق هذا القول على وجه الخصوص

بالنسبة للاختبارات الجماعية التي تطبقها على جماعات كبيرة وتصحبها الآلات الحاسبة الإلكترونية. أي أنه في هذه الحالات جميعاً لا يتطلب الأمر إلا التأكد من التتبع الدقيق لإجراءات التطبيق والتصحيح وتكون المشكلة هي الضبط التجريبي لظروف إجراء الاختبار.

إلا أننا نلاحظ أنه في الاختبارات الفردية مثل مستفرد بيليه يلعب للممتحن دوراً أكثر تعقيداً، لأن إجراءات الاختبارات في هذه الأحوال ليست معقدة تقنياً جامداً بل يعتمد معظمها على نجاح الممتحن في تكوين العلاقة بينه وبين المفحوص، وعلى دوافع المفحوص. وقد يتطلب الأمر أن يقوم الممتحن بتقويم أداء المفحوص خلال عملية تطبيق الاختبار لأن مثل هذا الأداء يحدد كيف يستمر الممتحن في موقف الاختبار. ولذلك فقد يختلف الممتحنون في النتائج التي يحصلون عليها من نفس المفحوصين حتى ولو توفرت في هؤلاء الممتحنين أفضل الشروط وهذا الاختلاف قد يكون من نوع تبليين للخطأ الذي يمكن أن ترجعه إلى الفروق الفردية بين الممتحنين ويؤثر بالطبع في التباين الكلي لدرجات المفحوصين. ولذلك فإذا كانت الاختبارات التي نستخدمها من النوع الذي تلعب فيه فردية الممتحن دوراً هاماً، أو إذا كانت أدوات التقويم من الذي يتأثر بالعوامل الذاتية مثل المقابلة الشخصية، نكون في حاجة ماسة إلى تقدير " ثبات الممتحن " وخاصة إذا كان من الضروري معرفة العلاقة بين تعديرات أكثر من ممتحن لنفس الفرد أو الأفراد.

وبالمثل نجد أن بعض الاختبارات تثير مشكلة " ثبات المصحح Reliability of Scorer وخاصة لاختبارات المقال والاختبارات الإسقاطية

بروجه عام، وهي الأدوات التي تسمح بمقدار كبير من التفسير الذاتي من جانب المصحح وهو للممتحن أيضاً في كثير من الأحوال.

ومن الواضح أن معامل ثبات المصحح يمكن للحصول عليه بحساب معامل الارتباط بين الدرجات التي يعطيها مصححان أو أكثر لنفس الاختبارات أو نفس الأفراد. بعبارة أخرى فإن كل مفحوص يحصل على درجتين أو أكثر من تصحيح اختبار واحد. وهذه الطريقة لا تتضمن بالطبع نواحي التقصير الموجودة في طريقة إعادة الاختبار. أما معامل ثبات الممتحن فهو أكثر صعوبة لأنه بطبيعته يتضمن يتضمن إعادة اختبار، وبالتالي فإن قيام أكثر ممتحن بإعادة الاختبار يعني الوقوع في جوانب للتقصير التي سبق الحديث عنها في طريقة إعادة الاختبار التي يقوم بها نفس الممتحنين. بالإضافة إلى أن معامل الارتباط للنتائج يعكس عوامل الاستقرار الوقي وثبات الممتحن معاً. ويمكن للجوء إلى نمط مختلف من التخطيط التجريبي يعتمد على تحليل للتباين لعزل تباين الخطأ الذي يمكن إرجاعه إلى الممتحنين. وعلى كل ف يمكن أن نعتبر كلاً من معامل ثبات المصحح ومعامل ثبات الممتحن من نوع واحد ونطلق عليه اسم معامل الموضوعية Coefficient of objectivity.

ويمكن أن نلخص أنواع معاملات الثبات وطرق تقديرها في الجدول التالي.

جدول رقم (1)

المعامل	طريقة تقديره	تباين الخطأ
معامل الاستقرار	إعادة الاختبار بنفس صورته ونفس الممتحنين في مناسبة مختلفة	التغيرات المؤقتة
معامل التكاليف	تطبيق الاختبار مع صورته متكافئة له في نفس الوقت	عينة محتوى الاختبار
معامل الاستقرار والتكاليف معاً	تطبيق الاختبار ثم تطبيق صورة متكافئة له في مناسبة مختلفة	التغيرات الزمنية عينة المحتوى
معامل الاتساق الداخلي	التجزئة النصفية	عدم تجانس نصفي الاختبار
معامل الاتساق الدخلى	تحليل المفردات باستخدام طريقة كيودر ريتشاردسون	التجانس الكلي لأسئلة الاختبار
معامل الموضوعية	تكرار المصحح وتكرار الممتحن	ذاتية الممتحن، التغيرات الزمنية

صدق الاختبارات

يتعلق موضوع صدق الاختبارات بما يقسمه الاختبار وإلى أي حد يلجج في قياسه، وهو موضوع لا يقتصر على عملية القياس وإنما قد يمتد إلى المنهج التجريبي بصفة عامة والذي يتطلب اختبار الفروض العلمية للتحقق من صحتها.

وإذا كان الصدق مسألة على درجة كبيرة من الأهمية في تقرير صلاحية أو فائدة أي فرض علمي أو في تحديد معنى مفهوم معين فإله في الواقع هو الذي يحدد قيمة الاختبار ومغزاه. ويجب أن نقرر منذ البداية أن عنوان الاختبار لا يدل بالضرورة على ما يقسمه الاختبار لأنه عادة ما يكون مجرد كلمة أو عبارة قصيرة تهدف إلى التعرف على الاختبار. ولا يمكن أن نستدل على الوظيفة السلوكية التي يقيسها الاختبار إلا بفحص مجموعة من المحكات (أو غيرها من المصادر للموضوعية للمعلومات) المستقلة نسبياً عن الاختبار.

ويمكن القول بوجه عام أن جميع طرق تحديد صدق الاختبارات تبحث عن العلاقة التي تربط أداء الأفراد في الاختبار بحقائق ومعلومات خاصة بالوظيفة السلوكية موضوع الاختبار يمكن ملاحظتها ملاحظة مستقلة. وهناك كثير من الطرق التي تستخدم في هذا الصدد ولها مسميات كثيرة، إلا أننا سوف نتحدث عن الأنواع الأربعة من الصدق التي وصفها كتيب التوصيات الفنية للجمعية الأمريكية لعلم النفس، والتي عالجتها أنستازي وكرونباك بالتفصيل.

أ- صدق المحتوى أو المضمون:

يقصد بصدق محتوى الاختبار أن نقوم بفحص مضمون الاختبار فحصاً دقيقاً منظماً لتحديد ما إذا كان يشمل على عينة ممثلة لميدان السلوك الذي يقيسه.

ومثل هذا النوع من الصدق نستخدمه في العادة في تقويم الاختبارات التحصيلية التي نعدّها لقياس مقدار ما تمكن التلاميذ من إتقانه من جوانب مهارة معينة أو ما حصلوه من مقرر دراسي أو برنامج مهني. وقد يبدو أن مجرد فحص مضمون الاختبار يكفي للتكليل على صفته وصلاحيته للغرض الذي وضع من أجله، فاختبار معين في الضرب أو التاريخ قد يبدو صادقاً بالتعريف إذا اشتمل على أسئلة تدور حول عملية الضرب أو التاريخ.

ولكن الأمر ليس على هذا القدر من السهولة لأن من أهم المشكلات التي نواجهها هنا هي مشكلة عينة المحتوى. فميدان المحتوى الذي نختبره يجب أن يتم تحليله تحليلًا دقيقًا منتظمًا حتى نتأكد من أن جميع الجوانب الأساسية أمكن أن تغطيها أسئلة الاختبار تغطية ماثمة وبالنسب الصحيحة حتى لا يكون لدينا عدد كبير من الأسئلة في جانب معين على حساب الجوانب الأخرى. وهذا يتطلب بداء ذي بدء وصفًا تفصيليًا دقيقًا للمحتوى أو المضمون الذي نعد له الاختبار التحصيلي، كما يتطلب شمول الاختبار لأهداف للتدريس أو للتدريب وليس مجرد المادة الدراسية أو المقررات. وبذلك يصبح صدق المحتوى أكثر اعتماداً على مدى الاتصال بين استجابات الفرد على أسئلة الاختبار وبين ميدان السلوك الذي نختبره.

وتوجد عدة طرق تجريبية للتحقق من صدق المحتوى. فمثلاً إذا أمكن إعداد صورتين متكافئتين من الاختبار فيمكن أن نطبق إحداهما قبل التدريب والأخرى بعده للتحقق من حدوث أي تحسن في الدرجات. وهناك طرق أخرى مثل دراسة أنماط الأخطاء الشائعة في الاختبار، وتحليل طرق العمل التي يستخدمها المفحوصون وذلك بإعطاء الاختبار فردياً مع توجيههم إلى " التفكير

بصوت مرتفع * لثناء حل مشكلات الاختبار. كما أن أثر السرعة يمكن تحديده بمعرفة عدد الذين فشلوا في الانتهاء من الاختبار أو بأي طريقة أخرى من الطرق الملائمة.

وعندما يكون الهدف أن نكتشف أثر القدرة على قراءة التعليمات في الأداء الاختباري (وخاصة إذا كانت تلك القدرة غير مرتبطة بهذا الأداء) فيمكن أن نحسب معامل الارتباط بين درجات الاختبار التجريبي ودرجات اختبار مقنن في فهم للقراءة. أما إذا كان الاختبار التجريبي يهدف إلى قياس القدرة على فهم للقراءة فيمكننا أن نعطي للأفراد الأسئلة التي يحتويها الاختبار دون أن نعطيهم النص للقرائي الذي تعتمد عليه لمعرفة عدد الأسئلة التي يمكنهم أن يجيبوا عليه معتمدين على معلوماتهم السابقة، إلى غير ذلك من العوامل التي قد تؤثر في مثل هذه الاختبارات.

وكما سبق أن قلنا أن صدق المحتوى يفيد في مراجعة اختبارات التحصيل ولكنه لا يكون كافياً في اختبار القدرات والاستعدادات ومقاييس الشخصية، بل قد يكون مضللاً في كثير من الأحيان. فعلى الرغم من أن هذه الاختبارات لا بد أن يتوفر فيها معك الملائمة والفاعلية بالنسبة لمضامينها وخاصة في المراحل الأولى لإعدادها، فإننا فيما بعد نلجأ إلى وسائل التحقيق التجريبي المضبوط التي سوف نتحدث عنها فيما بعد لتقدير صحتها. ويتضح ذلك إذا علمنا أن اختبارات القدرات والاستعدادات والشخصية على عكس الاختبارات التحصيلية لا تتشابه تشابهاً داخلياً كثيراً مع ميدان السلوك الذي تعد هذه الاختبارات عينة منه، وبذلك لا يكون لمضامينها أي دور في الكشف عن القروض التي أنت بمؤلف الاختبار

إلى انتقاء مضمون معين يستخدمه في قياس صفة معينة، لأن هذه الفروض تتطلب تحقيقاً تجريبياً حتى يمكن تحديد صدق الاختبار.

كما أن اختبارات القدرات والشخصية على عكس الاختبارات التحصيلية لا تعتمد على مقرر دراسي أو على برنامج تدريبي أو على مجموعة موحدة من الخبرات السابقة يمكن منها صياغة مضمون الاختبار، وإنما يتفاوت الأفراد كثيراً في طرق عملهم وفي العمليات النفسية التي يستخدمونها عند الإجابة على نفس الأسئلة، وعلى ذلك فإن الاختبار الواحد قد يقيس وظائف سلوكية مختلفة في الأفراد المختلفين. وفي مثل هذه الأحوال يستحيل علينا تحديد الوظائف السلوكية التي يقيسها الاختبار من مجرد فحص محتواه فقد يحل طلاب الجامعة مشكلة معينة مستخدماً للتعبير اللغوي (أو للرمزي عموماً ومنه الرياضي)، بينما يحل الميكانيكي نفس المشكلة ويصل إليه نفس الحل مستخدماً للتصور البصري المكاني. كما أن اختباراً معيناً للاستدلال الحسابي لطلاب المرحلة الإعدادية قد يقيس مجرد الفروق الفردية في سرعة العمليات الحدية عند طلاب الجامع.

ويجب ألا نخلط بين صدق المحتوى والصدق الظاهري. والواقع أن الصدق الظاهري ليس صديقاً بالمعنى العلمي للكلمة لأنه يدل على ما يبدو أن الاختبار يقيسه من الظاهر لا على ما يقيسه الاختبار بالفعل. كما أنه يعني ما إذا كان الاختبار يبدو صادقاً في نظر المفحوصين من ناحية، وفي نظر الفاحصين، سواء أعدوا إعداداً قديماً لم يعدوا، من ناحية أخرى. والواقع أن موضوع الصدق الظاهري يدخل في باب العلاقة الطيبة التي يكونها الفاحص مع المفحوص، وهو من الأمور المرغوب فيها نون شك رغم ما يسببه من خلط

وتدخل بينه وبين غيره من مفاهيم الصدق الحقيقية. فيما لا شك فيه أنه إذا بدأ الاختبار للفاحصين أو المفحوصين لا معنى له، أو لا علاقة له بالظاهرة موضوع القياس، أو غير مناسب أو سانجاً أو سخيلاً، فمن المضمّن أن تعاون هؤلاء سيكون ضعيفاً حتى ولو كان الصدق الحقيقي للاختبار عالياً. بعبارة أخرى لا بد أن يتوافر في الاختبار ما يسمى للصدق الظاهري حتى يكون أكثر فاعلية في مولف القياس العملية.

ب- الصدق التنبئي:

يقصد بالصدق التنبئي قدرة الاختبار وفاعليته في التنبؤ بنتيجة معينة في المستقبل، ولهذا يقوم على المقارنة بين درجات المفحوصين في الاختبار ودرجاتهم في مقياس آخر مباشر للأداء اللاحق للفرد وهذا المقياس الحديدي يسمى من الناحية الفنية " المحك أو الميزان. وهذا اللوح من الصدق يتصل اتصالاً وثيقاً بالاختبارات التي تستخدمها في أغراض الانتقاء والتصنيف والتصفية سواء في الأغراض التربوية أو المهنية أو العسكرية، وفي تحديد المرضى للعصابيين الذين يحتمل أن يفيدوا من برنامج علاجي معين. وفي جميع هذه الأغراض تتشابه إجراءات الصدق، حيث نعطي الاختبار لعينة من الأفراد تمثل للمجتمع الأصلي (موضع الاهتمام)، ولكن الدرجات التي يحصل عليها أفراد هذه العينة لا تستخدم في اتخاذ أية قرارات خاصة بهم. كما يجب ألا تعطى هذه الدرجات إلى أي شخص أو سلطة يمكن له أو لها أن تتخذ قرارات تؤثر فيهم، حتى لا يطرأ على المحك أي تشويه أو خلط. وبعد ذلك تتم متابعة العينة الأصلية حتى تتوافر لدينا عنها المعلومات التي نستقيها من المحك والتي تؤدي بنا إلى تحديد

المدى الذي تتفق فيه درجات الاختبار مع النتائج التي نلاحظها كما تتمثل في درجات المحك، وهذا الاتفاق يحدد لنا مدى تنبئية الاختبار على أسس إحصائية.

وعادة ما يتضمن دليل الاختبار قوماً خاصة بالصدق للتنبئ مشتقة من استخدام محكات مختلفة بهدف مساعدة من يستخدم الاختبار في فهم ما يقسمه وذلك على الرغم من أن مستخدمي الاختبارات قد لا يهتمون اهتماماً مباشراً بالتنبؤ بأي من المحكات المستخدمة. ومعظم المحكات التي تذكرها كتيبات الاختبارات عادة ما تكون للحصول المدرسي الأكاديمي العام، أو أداء الأفراد في برنامج تدريبي متخصص، أو أداء الأفراد في ميدان العمل أو التعلم نفسه. ونفصل الحديث عن هذه المحكات كما يلي:

1- للحصول المدرسي الأكاديمي العام:

هو أشهر هذه المحكات جميعاً وخاصة بالنسبة لاختبارات الذكاء. ولذلك توصف هذه الاختبارات عادة بأنها مقاييس الاستعداد المدرسي. أما المؤشرات التي نلجأ إليها كمقياس لمحك الحصول الأكاديمي فغالباً ما تكون للدرجات أو للتقديرات المدرسية، ودرجات الاختبارات التحصيلية، ونتائج النقل والتخرج، ودرجات الشرف والتخصص، والجوائز والملح، وتقديرات المعلمين لذكاء تلاميذهم، هذه المؤشرات المختلفة تمنطينا معلومات خاصة بهذا المحك في جميع المستويات التعليمية من المرحلة الابتدائية حتى الكليات الجامعية والمعاهد العليا. وهذا المحك نستخدمه كما قلنا في تقدير صدق اختبارات للذكاء، وكذلك بطاريات الاستعدادات للفرقة، واختبارات القدرات الطفولية والخاصة، وبعض لختبارات الشخصية. وعند حساب صدق الاختبار بفرض استخدامه في عمليات الانتقاء لأكاديمية معينة فمن الشائع أن نستخدم محكاً معيناً هو متوسط

درجات كل طالب مستجد. هذا المتوسط هو بدوره متوسط درجات الطالب في جميع المقررات التي يدرسها في السنة الأولى أو في السنة الإعدادية بالجامعة.

2- الأداء في برنامج تدريبي أو تعليمي معين:

وإذا كان ما نقوم به هو إعداد بطاريات الاستعدادات الفارقة والقدرات الطائفية فإن المحك الذي نستخدمه هو أداء الأفراد في برنامج معين من التدريب المتخصص. ففي حساب صدق اختبار معين للقدرة الميكانيكية يمكن الاعتماد على التحصيل النهائي في مقرر معين يتعلق بإحدى المهن للصناعة. وكذلك يمكن أن نعتبر التحصيل في مقررات التعليم التجاري مثل الاختزال والآلة الكاتبة وميك الدفاتر، محكات لاختبارات الاستعدادات في هذه الميادين. وكذلك فإن أداء طلاب المعاهد الموسيقية والفنية قد يعتبر محكاً لحساب صدق اختبارات الاستعدادات الموسيقية والفنية. وقد حسب صدق معظم اختبارات الاستعدادات المهنية باستخدام محكات التحصيل في كليات الحقوق والطب والصيدلة والهندسة وغيرها. أما إذا كانت الاختبارات من النوع الذي نستخدمه في أغراض برنامج معين للانتقاء. فقد نستخدم مجالات التدريب كمصدر للمعلومات التي نحتاج إليها، كما هو الحال في حساب صدق اختبارات انتقاء الطيارين حيث يكون المحك هو أداء الطيارين في مواقف التدريب الفعلية.

ومن المؤشرات الخاصة التي يتضمنها محك الأداء في برنامج التدريب يمكن أن نذكر الاختبارات التحصيلية التي نطبقها عقب الانتهاء من برنامج التدريب أو تدريبات المدربين أو النجاح والفضل في التدريب (الاستمرار والإكمال في مقابل الانقطاع والاستعدادات). وقد نستخدم درجات المدرسة الثانوية أو الكليات الجامعية والمعاهد العليا في مقررات معينة في تحديد صدق

بطاريات الاستعدادات، فمثلاً يمكن مقارنة درجات الطلاب في اختبار الفهم اللغز، بدرجاتهم في مقررات اللغة العربية، ودرجات التصور البصري المكاني بدرجات مقرر اللغة العربية، ودرجات التصور البصري المكاني بدرجات مقرر الهندسة... إلخ.

ولا بد أن نميز في هذا الصدد بين ما يسمى المحكات المتوسطة والمحكات القصوى أو النهائية بالنسبة لمؤشرات سجلات التدريب. فعند إعداد اختبار لقياس الاستعدادات للطيران أو لقياس الاستعداد الطبي مثلاً نجد أن المحكات النهائية هي الأداء الفعلي في ميدان القتال، أو النجاح في ممارسة الطب. ويتطلب ذلك دون شك وقتاً طويلاً حتى يمكننا الحصول على معلومات من هذه المحكات، بل أنه من المشكوك فيه إمكان وصولنا إلى محك نهائي في الواقع بالإضافة إلى أنه، حتى إذا توفر لنا مثل هذا المحك النهائي، فإنه قد يتعرض لعوامل كثيرة يصعب التحكم فيها مما يجعل المحك في النهاية لا فائدة منه. فمن الصعب مثلاً أن تقوم درجة النجاح النسبي للأطباء الذين يمارسون مختلف التخصصات في أجزاء مختلفة من الوطن ولهذا السبب تلجأ في العادة إلى المحكات المتوسطة التي ذكرنا أمثلة لها فيما سبق مثل سجلات الأداء في مرحلة معينة من مراحل التدريب.

3- الأداء على العمل نفسه:

هو أكثر المحكات صلاحية لأنه يعتمد على سجلات تتبعية للأداء الواقعي والفعلي للأفراد في العمل. وقد استخدم العلماء هذا المحك إلى حد ما في حساب صدق اختبارات الذكاء العام وللشخصية وإلى حد كبير في حساب صدق اختبارات الاستعدادات الفارقة والقدرات الطاقية. ومن المؤكد أن " الأعمال "

تختلف في مستوياتها وفي أنواعها. ويصدق هذا على الأعمال الصناعية والتجارية والمهنية والعسكرية وغيرها، ورغم أن معظم مقاييس أداء " العمل " قد لا تعتبر محكات نهائية إلا أنها أفضل بكثير من محك سجلات التدريب الذي تحدثنا عنه آنفاً. ومع ذلك فقد لا يسمح قياس أداء العمل بتوحيد وتقنين الشروط التي تتوفر أثناء التدريب، كما أنه يتطلب تتبعاً طويل المدى مما يؤدي إلى قلة عدد المفحوصين، وبالتالي يؤثر في النتائج على نحو لا حيلة للباحث فيه.

ج- الصدق التلازمي (أو المصاحب):

الصدق التلازمي أو المصاحب هو نوع الصدق الذي يدل على وجود علاقة بين درجات الاختبار ومؤشرات المحك التي نحصل عليها في نفس الوقت تقريباً.

ويلجأ الباحثون إلى الصدق التلازمي في بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبئي وما تتطلبه من وقت طويل، وفيها نعطي الاختبارات لمجموعة من الأفراد تتوافر لدينا عنهم معلومات خاصة بمحك معين، فنقارن مثلاً درجات طلاب الجامعة في اختبار معين بمتوسط تراكمي لدرجاتهم الدراسية حتى وقت الاختبار كما نقارن درجات العمال في الاختبار بنجاحهم الراهن في العمل مثلاً.

ويلجأ الباحثون إلى الصدق التلازمي في بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبئي وما تتطلبه من وقت طويل، وفيها نعطي الاختبارات لمجموعة من الأفراد تتوافر لدينا عنهم معلومات خاصة بمحك معين، فنقارن مثلاً درجات طلاب الجامعة في اختبار معين بمتوسط تراكمي

لدرجتهم للدراسية حتى وقت الاختبار كما نقارن درجات العمال في الاختبار بنجاحهم الراهن في العمل مثلاً.

وتتمثل أهم مشكلات الصدق التلازمي في صعوبة تحديد علاقات السبب والأثر، لأنه يدل بالفعل على علاقة القتران أكثر من دلالاته على علاقة "علية"، ولذلك فإن استخدامه كبديل للصدق التنبئي يتضمن صعوبات منطقية ظاهرة، لأن التنبؤ يتضمن في جوهره هذه العلاقة السببية.

ومن ناحية أخرى فإن للصدق التلازمي استخداماته الخاصة في أغراض معينة فهو أكثر ملائمة للاختبارات التشخيصية والتي تلجأ إليها في حل المشكلات الراهنة أكثر من مشكلات المستقبل. فعند إعداد مقياس للشخصية يهتم الصدق التلازمي بسؤال مثل: هل هذا الطفل عصبي ؟، أما الصدق التنبئي فيجيب على سؤال من نوع آخر وهو: هل من المحتمل أ، يصير هذا الطفل صعباً ؟ ومع ذلك فكثير ما تثار التساؤلات حول وظيفة الاختبار في حالة الصدق التلازمي الذي تتصاحب فيه معلومات الاختبار ومعلومات المحك والواقع أن الاختبار في مثل هذه الحالة عادة ما يكون بديل أبسط وأسرع وأقل تكلفة وجهداً من معلومات المحك. فإذا كان المحك يتطلب ملاحظة مباشرة لمرضى مدة أسبوعين من لحظة دخوله المستشفى، فإن الاختبار الصادق صدقاً تلازمياً يمكنه أن يميز بين الأسوياء والعصابيين والحالات المشكوك فيها في وقت قصير وبذلك يختصر لنا عدد هذه الأشخاص الذي يتطلبون هذه الملاحظة الشاملة، بالإضافة إلى اقتصاد الوقت والجهد الذي يبذل في الملاحظة المباشرة.

وقد يتضمن دليل الاختبار معلومات عن الصدق التلازمي إما على أنه بديل للصدق التنبئي (رغم صعوبات ذلك) أو كمؤشر القوة التشخيصية للاختبار.

وهذا النوع من الصندق لكثير ملاءمة دون شك لاختبارات الشخصية ومع ذلك فقد يستخدم في حساب صندق اختبارات الاستعدادات والقدرات.

1- المجموعات المتضادة:

ومن للمحكات الأساسية التي نلجأ إليها في حساب الصندق للتلامي ما يسمى " المجموعات المتضادة " وهو من أنواع المحكات المركبة التي تدل على الآثار المترتبة وغير المضبوطة للحياة اليومية، أي أنه لكل وضوحاً وأكثر تعقيداً من المحكات الأخرى التي سبق مناقشتها، ويعتمد في أساسه على بقاء الفرد واستمراره داخل مجموعة معينة أو عدم بقاءه فيها وخروجه منها. ففي حسابنا لصندق اختبار الذكاء يمكن أن نقارن درجات ضعاف العقول المقيمين بإحدى المؤسسات بالدرجات التي يحصل عليها تلاميذ مدرسة عادية في نفس الزمن تقريباً. في مثل هذه الحالة يتأثر المحك بالعوامل العديد التي تتحكم في هذه المؤسسات بوجه عام. وكذلك الشأن إذا أردنا حساب صندق اختبار الاستعداد للموسيقى أو الاستعداد الميكانيكي فيمكن أن نقارن بين الدرجات التي يحصل عليها الطلاب الذين يدرسون للموسيقى أو الهندسة في مدرستين أو معهدين متخصصين، وبين درجات طلاب آخرين يدرسون في مدرسة ثانوية عادية أو كلية جامعية لم يتم اختيارهم على أساس الوظيفة السلوكية موضوع القياس.

وتستخدم طريقة المجموعات المتضادة في حساب صندق اختبارات الشخصية على وجه الخصوص، فمثلاً يمكن أن نقارن أداء من يمارسون أعمالاً قيادية أو إقناعية بأداء من يمارسون أعمالاً كتابية في اختبار معين للسمات الاجتماعية ليتمكن حساب صندق هذا الاختبار، مفترضين أن الأفراد الذين يعملون بمهن معينة تتطلب للتعامل الاجتماعي (البيع، التدريس... إلخ) وظلوا

يقومون بها سوف يتقنمون على غيرهم من الأفراد الذين يعملون في ميادين أخرى لا تتطلب مثل هذا النوع من التعامل كالأعمال الكتابية والهندسية. بل يمكن أن نقارن بين أداء الطلاب الذين يقومون بنشاط كبير خارج المنهج وأداء غيرهم من الطلاب الذين لا يقومون بهذه الأنشطة خلال فترة دراسية معينة. كما أننا كثيراً ما نلجأ إلى الجماعات المهنية لحساب صدق اختبارات الميول (ومنها الميول المهنية) ومقاييس الاتجاهات. ومن الجماعات التي تستخدم في حساب صدق مقاييس الاتجاهات الجماعات السياسية أو الدينية أو غيرها من الجماعات الخاصة التي تمثل وجهات نظر مختلفة لاختلافاً واضحاً حول موضوعات اجتماعية معينة. وقد تستخدم جماعات معينة مثل الأحداث المنحرفين المدعين في المؤسسات في مقابل الأحداث غير المنحرفين، أو العصاة في مقابل الأسوياء، عند حساب صدق اختبارات الشخصية التي تتعلق بالتكيف الانفعالي أو الاجتماعي. كما أن في بعض هذه الاختبارات يلجأ الباحثون إلى التشخيص الطبي النفسي باعتباره أساساً لاختبار الأمثلة ودليلاً على صدق الاختبار، إلا أن التشخيص الطبي النفسي لا بد أن تتوفر فيه عدة شروط هامة ليصبح محكاً مفيداً، من أهمها أن يستند هذا التشخيص إلى ملاحظة طويلة المدى ودراسة مفصلة لتاريخ الحالة، أما إذا اعتمدنا على مجرد مقابلة شخصية قصيرة أو على فحص طبي نفسي غير عميق فإن يكون هناك سبب يدعونا إلى توقع أن يكون تشخيصنا أفضل من الدرجة التي يحصل عليها الفرد في الاختبار. نستدل من أي منهما على الحالة الانفعالية للعميل. بعبارة أخرى يصبح التشخيص الطبي النفسي كالاختبار مجرد مؤشر أو مئبى يحتاج في ذاته إلى حساب صدقه ولا يكون محكاً نعتمد عليه في تقدير صدق الاختبارات الأخرى.

2- التلذيرات:

وقد نلجأ في حساب الصلوق للتلازمي لاختبارات للشخصية إلى التلذيرات Ratings. وقد سبق أن تحدثنا عن التلذيرات التي يعطيها المعلمون لتلاميذهم أو المشرفون على العمل لعمالهم وغير ذلك علد حديثنا عن المحكات الأخرى والتي نطليها بهلف الحصول على معلومات تتصل بهذه المحكات مثل التلصصول للمدرسي الأكليمي أو الأداء في نوع متخلص من الأعمال أو اللجاح المهني علد حساب التلذيري. إلا أن التلذيرات قد تصبح في ذاتها محكاً لعلد عليه في حساب الصلوق للتلازمي وخاصة إذا لم تقتصر على جانب معينة كالتلصصول مثلاً ولكن تتلصصن حكماً متخلصاً من جانب الملاحظ (أو الفاحص) على إحدى الصلغات التي ترمع الاختبارات النصلبة قياسها. وإذلك فقد نطلب من الفاحصين أن يعطوا تلذيرات للمفحوصين في خصائص معينة مثل السيطرة أو التقيادة أو الأصالة أو الابتكار أو الأملة... إلخ. وهذه التلذيرات التي في ذاتها تعد محكات لستلخدمها الباحثون في حسابهم لصلوق كثير من الاختبارات، وبخاصة لختبارات الشخصية كما سبق أن قلنا، وإذلك لصعوبة الحصول على محكات موضوعية في هذا الميدان وبخاصة بالنصلبة للصلغات الاجتماعية التي لا يمكن للحكم عليها إلا عن طريق التلذيرات المعتمدة على العلاقة للشخصية.

والمواقع أن التلذيرات التي نلصل عليها من لشخلص لديم التلذيري والخبرة وفي ظروف مضبوطة ضبطاً تجريبياً دققاً نوفر لنا معلومات هامة عن المحك ومما لا شكل فيه أن زيادة نلكتنا في هذه التلذيرات تتلطلب للحصول على تلذيرات مستقلة من أكثر من فاحص حتى يمكن التلطلب على العوامل للذاتية التي تؤثر في ثبات هذه التلذيرات والتي سبق لنا مناقشتها في حديثنا عن ثبات

الاختبارات. كما يمكن أن تردّد دقة التقديرات إذا استخدمنا مقاييس التقدير rating scales بحيث تتضمن وحدات محدّدة تحديداً واضحاً مع وجود ضمانات خاصة للتخلص من أخطاء التقدير الشائعة مثل أثر أو وقع الهالة Halo Effect (وهي عبارة عن ميل للمدّرين أو الفاحصين إلى للتأثر بسمة واحدة مطلوبة، وتكون في صلتح المفحوص أو في غير صلتحه والتي تؤدي بالتالي إلى للتأثير في أحكامهم الخاصة بالمسمات الأخرى)، كما لا بد أن تتوفر للمدّرين " لفية بالمسمات موضوع القياس " بحيث يمكنهم الحكم على وجودها بدرجة معينة لدى الأفراد الذين نختبرهم. بعبارة أخرى لا يكفي أن يعرف الفاحص للمفحوص فترة طويلة ليصبح قادراً على إصدار حكم عليه وإنما يجب أن تكون قد تهيأت له فرصة ملاحظته في مواقف مختلفة تظهر فيها السمة السلوكية موضوع الاهتمام. ولا يجب أن يعطي الفاحصون تقديرات عن المفحوص في سمة ليس لديهم عنها معرفة كافية.

وتلعب التقديرات دوراً هاماً في مقاييس المحك الذي يعتمد على التقويم الإكلينيكي. وفي مثل هذه الأحوال لا بد من تدعيم الحكم الإكلينيكي بعض المعلومات الإضافية مثل للتاريخ المفصل للحالة ودرجات الاختبارات وغيرها من للوسائل الموضوعية، أما للتقويم النهائي فيعتمد على عملية الحكم كما هو الحال في جميع التقديرات.

3- معاملات الارتباط بين الاختبار وغيره من الاختبارات:

ومن للوسائل الأخرى التي يستخدمها الباحثون في تقدير الصدق التلازمي حساب معاملات الارتباط بين درجات الاختبار التجريبي (الجديد) وبين درجات اختبارات أخرى قائمة بالفعل. وقد نلجأ إلى هذه الطريقة إن كان الاختبار الجديد

صورة مختصرة أو مبسطة لاختبار موجود فعلاً وفي هذه الحالة يكون الاختبار الأصلي مقياساً للمحك. كما يمكننا أن نقرر صدق اختبار لفظي (ورقة وقلم) باستخدام اختبار يدوي عملي لا يستغرق وقتاً طويلاً وله معامل صدق مرتفع ثم تقديره من قبل. وكذلك يمكن تقدير صدق اختبار جماعي باستخدام اختبار فردي (مثل اختبار ستانفورد بينيه) الذي استخدم كثيراً كمحك لحساب صدق الاختبارات الجماعية للنكاه). وفي مثل هذه الأحوال يمكن اعتبار الاختبار الجديد أحسن الحالات صورة تقريبية للاختبار الأصلي. ويجب أن نلاحظ أنه ما لم يكن الاختبار الجديد بديلاً مبسطاً أو مختصراً للاختبار الأصلي لا يمكن أن نقوم باستخدام هذه الطريقة وإلا كانت نتائجنا موضع شك شديد.

صلى التكوين الفرضي:

يصدق يقصد للتكوين الفرضي مدى قياس الاختبار لتكوين فرضي معين أو سمة معينة. ومن أمثلة هذه التكوينات للفرضية للنكاه والفهم الميكانيكي والطلاقة وسرعة المشي والعصابية والقلق... إلخ. ويعتمد هذا النوع من الصدق على وصف أوسع ويتطلب معلومات أكثر حول السمة السلوكية موضوع للقياس لحصل عليها من مصادر مختلفة.

وهناك أساليب كثيرة يمكن أن نلجأ إليها للحصول على أكبر قدر من المعلومات التي تلقي الأضواء على طبيعة السمة التي نهتم بها. ومن هذه الطرق ما يأتي:

1- تمايز العمر:

يمكن أن نلجأ إلى هذه الطريقة في حساب صدق اختبارات الذكاء (كاختبار ستانفورد بينيه، أو معظم اختبارات الذكاء لأطفال ما قبل المدرسة)، وذلك لتحديد ما إذا كانت الدرجات تتزايد بتقدم العمر مفترضين أن القدرات تنمو بتقدير العمر خلال الطفولة وبالتالي نتوقع زيادة في درجات الاختبار إذا كان الاختبار صادقاً. ويعود الفضل إلى بيليه في صياغة هذا المفهوم الخالص بتحديد للقياس العمري للذكاء.

ومن الواضح أن محك تمايز العمر لا يمكن استخدامه في الوظائف السلوكية التي لا تظهر تغيرات مطردة مع تقدم العمر ومن ذلك مقاييس الشخصية بوجه عام. كما أننا يجب أن نشير إلى أن هذا المحك محك سلبي وليس محاً إيجابياً. بعبارة أخرى إذا لم تستطع درجات الاختبار أن تتزايد مع تزايد العمر فإن ذلك قد يدل على أن الاختبار غير صادق في قياسه للقدرات التي يدعي أنه عيلة لها، إلا أن تحسين الاختبار بحيث يصبح مقياساً متزايداً مع تزايد العمر لا يحدد لنا في ذاته الميدان الذي يقيسه الاختبار. فمثلاً يظهر كل من الطول والوزن زيادة منتظمة مع تزايد العمر ومع ذلك لا يمكن أن نعتبر أحدهما أو كليهما من مقاييس الذكاء. وكذلك يجب أن نؤكد أنه في الاختبار النفسي الذي يقيس خصائص سلوكية تتزايد مع العمر إنما تتأثر هذه الزيادة أو النمو بظروف البيئة التي يقطن فيها الاختبار، فالظروف البيئية والحضارية قد تستثير بعض الخصائص السلوكية وقد تعطل البعض الآخر، ومعنى ذلك أن محك تمايز العمر لا يمكن أن نعتبره محكاً شمولياً بحال من الأحوال، وإنما هو كغيرها من المحكات محكوم بالموقف الحضاري للخاص الذي نشته منه.

2- معاملات الارتباط بالاختبارات الأخرى:

وقد لنجأ في تقدير صدق التكوين الفرضي إلى حساب معاملات الارتباط بين الاختبار التجريبي وبين الاختبارات التي تتشابه معه في الوظائف التي يقيسها والتي ثبتت صحتها، ويختلف محك حساب معاملات ارتباط الاختبار بالاختبارات الأخرى في صدق التكوين الفرضي عنه في الصدق للتلازمي في أن معاملات الارتباط في صدق التكوين الفرضي يجب ألا تكون مرتفعة ارتفاعاً كبيراً جداً كما هو الحال في الصدق للتلازمي، لأن الاختبار التجريبي الذي يرتبط ارتباطاً عالياً جداً باختبار جاهز دون أن تتوفر فيه أية مميزات إضافية يعني أن الاختبار التجريبي الجديد ما هو إلا لزوج وتكرار للاختبار السابق لا مبرر له.

ويمكن استخدام محك معاملات الارتباط بالاختبارات الأخرى بطريقة أخرى نحدد بها مدى تحرر الاختبار التجريبي من أثر العوامل التي لا علاقة بالوظيفة السلوكية موضوع القياس، فمثلاً إذا حصلنا على معامل ارتباط صفري بين اختبار معين للاستعدادات أو السمات الشخصية، وبين اختبارات الذكاء العام أو الاستعداد المدرسي واختبارات الفهم اللغوي فإن ذلك قد يعني أن هذه العوامل لا ترتبط بالاستعداد أو السمة موضوع القياس. ولذلك فكثيراً ما تتضمن كتيبات الاختبارات ومعاملات الارتباط باختبارات الذكاء العام والقراءة والفهم اللغوي على أنها مؤشرات غير مباشرة أو سلبية للصدق. ومن ذلك أنه إذا كانت هذه المعاملات مرتفعة فيكون ذلك مدعاة للشك في الاختبار، إلا أن معاملات الارتباط المنخفضة لا تعني في ذاتها أن الاختبار التجريبي صادق. ومهما يكن

الأمر فإن محك استخدام معاملات الارتباط في تقدير صدق التكوين الفرضي
أقرب إلى طريقة تقدير الصدق التلازمي التي سبق الحديث عنها آنفاً.

3- التحليل العملي:

وربما يكون التحليل العملي من أهم الوسائل التي تستخدمها في تقدير
صدق التكوين الفرضي. والتحليل العملي (4) هو أسلوب إحصائي لتحديد
السمات النفسية، وهو في أساسه يتضمن تحليل العلاقات بين البيانات كما تتمثل
هذه العلاقة في صورة معاملات الارتباط. نقطة البدء في التحليل العملي
حساب معاملات الارتباط بين كل اختبار والاختبارات الأخرى. وعندما نفحص
الجدول الذي يتضمن هذه المعاملات (والذي يسمى هنا مصفوفة الارتباط) قد
نلاحظ وجود تجمعات معينة بين الاختبارات مما يوحي بوجود سمات مشتركة
بينها. فمثلاً إذا كانت الاختبارات التي ترتبط بعضها ببعض ارتباطاً عالياً من
نوع اختبارات المفردات والمشابهات والأضداد وتكملة الجمل وكانت معاملات
ارتباط هذه الاختبارات بالاختبارات الأخرى منخفضة نستنتج من ذلك مبدئياً
وجود عامل الفهم اللغوي. إلا أن هذه التحليل الذي يعتمد على مجرد فحص
مصفوفة الارتباط عملية صعبة ومعقدة وغير مؤكدة بل مستحيلة وخالصة إذا
كانت بطارية الاختبارات كبيرة، ولذلك نستخدم الأساليب الإحصائية الأكثر دقة
لتحديد العوامل المشتركة التي يمكن اعتبارها مسؤولة عن معاملات الارتباط
النتيجة. ولا مجال للإفاضة في شرح هذه الطريقة الإحصائية للرفعة في هذا
المساق.

وفي التحليل العملي نلاحظ أن عدد المتغيرات التي نصف في ضوءها
أداء كل فرد تختصر باستمرار من عدد كبير من الاختبارات الأصلية إلى عدد

أصغر نسبياً من العوامل أو السمات المشتركة. ثم يمكن وصف الفرد في ضوء درجاته في العوامل القليلة المحدد. بعبارة أخرى يمكن القول أن أهم أغراض التحليل للعاملية تبسيط وصف للسلوك باختزال عدد المتغيرات، أي أنه يسير في اتجاه اللطم للصحيح من التعدد والكثرة التي تتمثل في متغيرات الاختبارات إلى قلّة التي تتمثل في العوامل المشتركة أو السمات.

4- صدق المفردات:

وهناك طريقة أخرى لتقدير صدق التكوين للفرضي تتمثل في حساب صدق المفردات التي يتكلف منها للمقاس. وقد دلجأ إلى هذه الطريقة على وجه الخصوص عند تقدير صدق اختبارات الشخصية. ويتلخص هذه الطريقة في اعتبار الدرجة الكلية في الاختبار نفسه محك للصدق. وقد نستخدم نوعاً من طريقة للمجموعات المتضادة تتمثل في انتقاء جماعات متطرفة على أساس الدرجات الكلية في الاختبار نفسه محك للصدق. وقد نستخدم نوعاً من طريقة المجموعات المتضادة تتمثل في انتقاء جماعات متطرفة على أساس للدرجات الكلية في الاختبار ثم نقارن أداء المجموعة العليا بأداء المجموعة الدنيا في كل سؤال من أسئلة الاختبار. فإذا فشل السؤال في إعطاء نسبة " نجاح " من أفراد المجموعة العليا أعلى من أفراد المجموعة الدنيا نعتبره سؤالاً " غير صادق " وبالتالي يجب حذفه أو تعديله. وقد دلجأ إلى أسلوب معاملات الارتباط لهذا الغرض ذاته والذي يتمثل في حساب معامل الارتباط الرباعي بين " النجاح والفشل " في كل سؤال من ناحية وبين الدرجة الكلية في الاختبار من ناحية أخرى، ولا يمكن أن نحفظ في الاختبار إلا بالأسئلة التي تحطي معاملات ارتباط دالة بين السؤال والاختبار. وإذا اتبنا هذه الطريقة في إعداد أسئلة

الاختبار فإن الاختبار ككل يمكن وصفه بأنه على درجة عالية من التماسق الداخلي حيث يميز السؤال في نفس اتجاه تمييز الاختبار. وقد نلجأ إلى تحويل هذه الطريقة بحساب معاملات الارتباط بين درجات الاختبارات الفرعية وبين الدرجة الكلية في الاختبار. فمعظم اختبارات الذكاء مثلاً تتكون من اختبارات فرعية عديدة تطبق منفصلة ثم نحصل على الدرجة الكلية للاختبار. وفي مثل هذه الاختبارات لا بد أن ترتبط درجات الاختبارات للفرعية بالدرجة الكلية، وتتل هذه المعاملات على الاتساق الداخلي في الاختبار ككل.

ومن الواضح أن معاملات التماسق الداخلي هذه سواء اعتمدت على أسئلة الاختبار أو على الاختبارات الفرعية إنما هي مقاييس للتجانس بالمعنى الذي أشرنا إليه. ومن المؤكد أن درجة تجانس أو تماثل الاختبار تتشابه إلى حد ما مع صدق التكوين للفرضي لأنها تساعد دون شك في تحديد ميدان السلوك أو السمة التي يقيسها الاختبار والتي يعد عينة لها. ومع ذلك فإن التماسق الداخلي في حد ذاته ليس دليلاً كافياً على الصدق لأنه إذا لم تتوفر لدينا وقائع خارجية ومستقلة عن الاختبار نفسه لا يمكن أن نعرف كثيراً على ما يقيسه الاختبار.

5- الإجراءات التجريبية:

ومن الوسائل الأخرى التي يمكن أن نلجأ إليها لتقدير صدق التكوين الفرضي ما تعطيه لنا الإجراءات التجريبية من معلومات حول أثر متغيرات معينة على درجات الاختبار. فإذا كان الاختبار التجريبي (الحديد) معداً لقياس القابلية للقلق مثلاً يمكن أن نطبق هذا الاختبار على مجموعة من الأفراد نوضع في موقف يثير القلق (مثل الامتحان في ظروف ضاغطة) ثم نحسب معامل الارتباط بين درجات الأفراد في اختبار القلق قبل التجريب وبين مختلف

المؤشرات الفسيولوجية وغيرها التي تدل على القلق خلال التجريب (موقف الامتحان) وبعده. ويمكن القول أن هذا الإجراء التجريبي يشبه إلى حد كبير ما نلجأ إليه في تقرير الصدق للتنبئي، باستثناء الدور الإيجابي للباحث في تقديم مثير للقلق والتحكم فيه. أما المقاييس المباشرة للدالة على القلق فيمكن اعتبارها محكاً نراجع عليه تنبؤاتنا من درجات الاختبار.

موضوعية التصحيح:

بعد تصحيح الاختبارات بإعطاء الفرد درجة أو تقديرًا، وتفسير هذه الدرجة هو خطوة هامة، ولو أنه في ذاته يعد مقبلة لاتخاذ قرارا عملي أو تفسير عملي عن الفرد أو مجموعة الأفراد موضوع للقياس. وعلى كل فإن مشكلات تصحيح الاختبارات لا تقل أهمية وخطورة عن المشكلات الأخرى التي نلجأ إليها في هذا الكتاب، وإن كانت أكثر وضوحاً في نمط الاختبارات الذي نسميه " امتحانات المقال ". ففي هذا النوع كثيراً ما يتعامل الشخص لماذا حصل على درجة منخفضة. ومن المؤكد أن السبب الرئيسي هو أنه في هذا النوع من الامتحانات يكون من الصعب تحديد الإجابة الجيدة وبالتالي يكون من الصعب إعطاء درجة مناسبة لاستجابة ليست صحيحة صحة كاملة.

ولا يمكن للباحث أن يجري دراسة الملوك كما لا يمكن أن يتخذ قراراً عملياً إذا لم تكن معايير التصحيح على درجة من الموضوعية، ولذلك يقترح البعض لطرح هذه المشكلة إما تحديد قواعد الحكم يلتزم بها المصححون أو استخدام الأسئلة الموضوعية التي تتطلب التعرف على الإجابة الصحيحة من بين عدة إجابات.

تصحيح الأسئلة التي تتطلب استجابات حرة:

لا يزال للقياس الفردي يستخدم أسئلة ومشكلات تتطلب درجة من الحكم أثناء التصحيح، ومع ذلك فمن الممكن إعداد بعض الوسائل التي تتوفر فيها درجة من الموضوعية في هذا التصحيح وبخاصة فيما يتعلق بالخصائص الرئيسية للسلوك. فعند تقويم سلوك معين مثل الخطأ أمكن للباحثين إعداد نموذج للتصحيح يحتوي على عيّنات من الخطوط تمثل مستويات مختلفة من الجودة وعلى المدرس أن يحدد العينة الأكثر شبهاً بخط التلميذ حتى يحدد الدرجة التي يعطيها له. وكذلك يمكن إعداد مقاييس مؤسسة على تقدير الإنتاج للحكم على نوع بعض ألوان السلوك الصناعي أو التجاري أو الإداري. ومع النجاح النسبي للوسائل الموضوعية في تصحيح بعض جوانب السلوك في الاختبارات التي تحتوي على استجابات حرة، إلا أن هذه الوسائل لم تتجح نجاحاً مؤكداً في تصحيح الاختبارات اللغوية، ومع ذلك فإن التباين بين المصححين يمكن تقليله إذا أعدنا نماذج توضح نمط التصحيح المتفق عليه بالنسبة لإجابات نموذجية. ومن أمثلة ذلك نماذج التصحيح الخاصة باختبار ستانفورد بينيه للذكاء، واختبار رورشاخ للشخصية واختبارات الابتكار الحديثة.

وعلى الرغم من أن نماذج التصحيح قد تكون مناسبة لأغراض القياس، إلا أننا نؤكد بعض جوانب الحذر حول تصحيح اختبارات الاستجابات الحرة بوجه عام. ومعظم مصادر الخطأ تنشأ في الواقع مما يمكن تسميته "المعالجة الشخصية" للمختبرين أو المصححين، فقد لوحظ مثلاً أن بعضهم يعطي درجات مرتفعة بصفة مستمرة بينما يميل البعض الآخر إلى إعطاء درجات منخفضة. كما أن الدرجة التي يعطيها المختبر أو المصحح تتأثر باستعداده العقلي وتحيزه

وتخمينه وأغراضه الشخصية وأهوائه، بل تتأثر بأثر الهالة التي تحيط بالمفحوص.

ويمكن للتقليل من أثر هذه العوامل الذاتية إذا لجأنا إلى ما يسمى بالتصحيح ' الأصمى ' أي يقوم المختبر أو المصحح بتقدير الدرجات للأفراد دون أن يعلم نتائج الاختبارات السابقة التي أجريت عليهم أو بالفئة التي ينتمي إليها في البحث (المجموعة التجريبية أو الضابطة) وبدون هذا الاحتياط يصعب علينا أن نتقاضي تأثير العوامل الذاتية في للدرجة المعطاة.

تصحيح الأسئلة التي تتطلب الاختبار من متعدد:

يتكون نموذج لتصحيح لما نسميه اختبارات التعرف من قائمة الإجابات الصحيحة أو مفتاح التصحيح يمكن أن يستخدمه أي شخص. وقد اخترعت عدة وسائل أكثر كفاءة في الحصول على الاستجابات وتصحيحها، ومنها النسخ الكربونية، وإعداد أوراق مستقلة للإجابة. ولا شك أن استخدام أوراق مستقلة للإجابة يؤدي إلى اختصار التكاليف، لأن كتيب الاختبار يمكن استخدامه أكثر من مرة، كما أن الإجابات يمكن تصحيحها بسهولة إما بمفتاح منقوب أو بواسطة الآلات الحديثة. أما في النسخ الكربونية فإن الوجه الخلفي لصفحة الأسئلة (أو ورقة الإجابة) تلتصق به وتخفيه ورقة كربونية، وبالطبع يطبع مفتاح الإجابة على هذا الوجه الخلفي. ويختار المفحوص الإجابة ثم يضع العلامة التي تثل على هذا الاختبار على الوجه الظاهر لورقة الأسئلة (أو الإجابة)، وبالطبع فإن هذه العلامة تطبعها ورقة الكربون على الوجه الخلفي للورقة، وعند التصحيح يقوم المختبر أو المصحح بكشف الوجه الخلفي الذي توجد فيه مربعات مطبوعة توضح بأن يجب أن تكون الإجابات الصحيحة (أي مفتاح الإجابة)، وبالطبع

يكون من أيسر الأمور جمع عدد العلامات الكربونية التي طبعت في المربعات الصحيحة.

وقد ازداد في الوقت الحاضر الاعتماد على الآلات الإلكترونية في تصحيح الاختبارات. وفي هذه الحالة يقوم المفحوص بتسويد مكان الإجابة بالقلم الرصاص، وتحتوي الآلة على أصابع مكمية توضح علامات للقلم الرصاص، لأن الجرافيت الموجود في هذه العلامات يوصل التيار الكهربائي. كما تحتوي الآلة على عداد يوضح للعدد الكلي للعلامات الموضوعة في أماكنها الصحيحة، كما أن الآلة يمكنها أن تسجل عدد الأخطاء، والدرجة التي تعتمد على (ص - خ)، وغير ذلك من أعلام للدرجات وبخاصة ما يتعلق بتصحيح أثر التخمين. ويمكن لبعض هذه الآلات أن يصحح للمئات من أوراق الإجابة تصحيحاً دقيقاً في الساعة. وقد تزايد عدد الأوراق الذي يمكن تصحيحه باستخدام الآلات الإلكترونية الحديثة.

$$\frac{خ}{ص - 1} \quad \text{أشهر معادلة لتصحيح أثر التخمين هي} \quad د = ص - \frac{خ}{1}$$

حيث تدل (ص) على الإجابات الصحيحة.

(خ) على عدد الإجابات الخاطئة.

(ن) على عدد الاحتمالات الاختبارات للإجابة عن السؤال للواحد.

والصعوبة الرئيسية التي تواجهها آلات تصحيح الاختبارات، والتي يهتم بها المختصون في تطبيق الاختبارات، هي أن هذه الآلات لا يمكنها أن تصحح تصحيحاً دقيقاً ما لم تكن الأماكن المخصصة للإجابة قد قام المفحوص بتسويدها

تسويداً كاملاً ونظيفاً وحسب المواصفات المطلوبة. فمن المعروف أن الآلة تعطي وحدة الدرجة عندما تظهر العلامة التي وضعها المفحوص في المكان الصحيح ومع ذلك فنحن نتوقع أن تكون العلامة التي وضعها المفحوص بالقلم الخفيف أو الثقيل، أو تكون عريضة أو ضيقة. ولذلك فلا بد أن يضع المفحوص هذه العلامة بقلم معين وأن يملأ المسافة المخصصة للإجابة في ورقة الإجابة ملاً كاملاً حتى نتأكد أن الآلة سوف تعدّها ولا تتجاهلها. بعبارة أخرى قد لا نتجاهل الآلة العلامات غير الكاملة والمحو غير النظيف للأخطاء. وقد يكون عدد الإجابات في ورقة الإجابة التي لا تتوفر فيها المواصفات اللازمة كثيراً، ولذلك نجد في بعض المؤسسات موظفين يقومون بفحص كل ورقة قبل أن تسلم للآلة لتصحيحها، وهؤلاء يقومون بمراجعة إجابات المفحوصين لا بفرض تصحيحها ولكن بفرض التأكد من أن طريقة الإجابة مناسبة للتصحيح، ومعنى ذلك زيادة تكاليف التصحيح. ولا شك أنه يمكن التغلب على هذه الصعوبة بإجراء الاختبار وتطبيقه على نحو سليم.

المعايير

الدرجات الخام:

تعطينا معظم الاختبارات وصفاً كمياً مباشراً لأداء الشخص نطلق عليه اسم الدرجة الخام. وقد تكون الدرجة الخام عدد الأسئلة التي أجاب عليها المفحوص إجابة صحيحة أو الزمن الذي استغرقه في الأداء.

ويفسر معظم الأفراد الدرجة الخام ويستخدمونها دون معرفة لجوانب القصور فيها لأن هذه الدرجات تتميز بالسهولة في الحصول عليها وبمعرفة لها من خبراتنا الطويلة الامتحانات للمدرسية.

وفي الواقع يمكن للقول أن الدرجة الخام في الاختبار النفسي هي في حد ذاتها لا معنى لها وليس لها أية دلالة، ولا يمكن أن تنص إلا بمقارنتها بمعيار معين. ومن المهم أن نؤكد في هذا الصدد أننا لا نستطيع تفسير الدرجات في الاختبار النفسي كما نفعل بالقيم التي نحصل عليها من مقياس للظواهر الطبيعية (الطول مثلاً). فمقاييس الظواهر الطبيعية لها صفر حقيقي وتحتوي على وحدات متسوية بتتابع المقياس، ولذلك فإن هذه المقاييس تسمح لنا بالمقارنة من الدرجات أو القيم التي نحصل عليها من المقياس مباشرة، فنقول مثلاً إن طول الطفل (س) ضعف طول الطفل (ص)، أما في المقياس للنفسى لا نستطيع تقرير مثل هذا القول. فلنفرض مثلاً أن الطفل (س) حصل على درجة مقدارها 10% من أسئلة اختبار في معاني الكلمات فهل يعني ذلك أن ذلك للطفل لا يعرف (إلا عثر الكلمات التي يجب أن يعرفها ؟ الإجابة على ذلك بالنفي لأن للفاحص قد لا يكون مهتماً إلا بالكلمات الصعبة غير المألوفة للطفل. وحتى لو حصل الطفل على الدرجة صفر في اختبار من هذا القبيل فإن ذلك لا يعني أن للطفل ليست لديه قدرة لغوية على الإطلاق. فالفرق بين الطفل (س) الذي يحصل على الدرجة صفر في الاختبار والطفل (ص) الذي يحصل على النهاية القصوى، أي يجيب على أسئلة الاختبار جميعاً، هو فرق في القدرة على معرفة عدد محدود من الكلمات يتكون منها الاختبار من بين آلاف الكلمات.

معنى ذلك أن الفرق بين الأفراد في الدرجات الخام لا يدل بالضرورة على وجود مسافات " حقيقية " بين الأفراد. فمثلاً في اختبار الذكاء قد يحصل التلميذ (أ) على الدرجة 53 والتلميذ (ب) على الدرجة 56 والتلميذ (ج) على الدرجة 59، أي تكون الفروق في الدرجات الخام فروقاً متساوية. فهل يعني ذلك أن الفرق بين (ج) و (ب) يساوي تماماً الفرق بين (ب) و (أ) ؟ لا يمكن الإجابة على مثل هذا السؤال إجابة مؤكدة، لأن الفروق بين الدرجات إنما تعتمد على الأسئلة (أو الوحدات) التي نستخدمها.

والطريقة الوحيدة الممكنة التي يمكن بها أن نتحدث عن " الفروق المتساوية " حديثاً له معنى تكون باستخدام محك عملي يحتوي على مستويات للقيم. واختلاف المستويات يؤدي إلى قيم عددية مختلفة لنفس الاختبار، فالتلاميذ الثلاثة السابق ذكرهم في مثالنا حصلوا على درجات خام تتساوى للفروق بينها في اختبار الذكاء، بينما نجد أن احتمال نجاحهم في مادة الهندسة مثلاً هو 0.70 ، 0.90 ، 0.96 ، على التوالي، بينما يكون احتمال نجاحهم في مهنة هندسية هو 0.0001 ، 0.50 ، 0.50 على التوالي. بعبارة أخرى تصير " للمسافات المتساوية " في أحد هذه المقاييس " مسافات غير متساوية " في مقياس آخر.

وعند تصير درجات الاختبار يكون أمام الفاحص أن يختار أحد الاحتمالات الآتية:

- 1- أن يقارن الدرجة الخام مقارنة مباشرة بمستوى معين للأداء، كما هو الحال في نظام للنهايات الصغرى في الامتحانات المدرسية.
- 2- أن يقارن الدرجة الخام بالدرجات الأخرى التي حصل عليها أفراد المجموعة التي اختبرها.

3- أن يقارن الدرجة الخام بدرجات مجموعة مرجعية تسمى عينة التقنين بواسطة جدول المعايير Norms الذي يتم إعداده كخطوة هامة من خطوات بناء الاختبار.

4- أن يستخدم جدول التوقع لتقدير الأداء الممكن للفرد في المستقبل.

ومن بين هذه الطرق نجد أن الطريقة للثلاثة أي طريقة مقارنة درجة الفرد بأداء عينة التقنين هي أكثرها شيوعاً. وعادة ما يرجع المصحح أو المختبر إلى جدول للمعايير الخاص بالاختبار ليعرف المدى العادي للأداء، ثم يحول الدرجة الخام إلى نوع معين من الدرجة المشتقة والتي تعتبر سجلاً دائماً للوضع النسبي للفرد. وأشهر أنواع الدرجات المشتقة هي المئينيات والدرجات المعيارية والأعمار العقلية ونسب الذكاء.

المئينيات .

أبسط وسائل المقارنة هي أن نرتب الدرجات التي يحصل عليها الأفراد إما ترتيباً تنازلياً أو تصاعدياً. وبذلك يمكن أن نحدد الوضع النسبي للفرد بالنسبة لأقرانه. ويعتمد الترتيب دون شك على عدد الأفراد في الجماعة. فإذا أردنا أن نحدد التغير في وضع الفرد من مناسبة معينة لأخرى ننشأ في هذه الحالة صعوبات خاصة باختلاف حجم المجموعات. ولتجنب هذه الصعوبات نتحول الرتب إلى درجات مئينية (وأحياناً تسمى رتب مئينية). والدرجة المئينية هي رتبة صورة نسبة مئوية. والدرجة المئينية للشخص تدل على نسبة الأفراد من بين المجموعة التي تقع في مرتبة أقل منه. فإذا افترضنا أن لدينا 40 شخصاً منهم 27 شخصاً حصلوا على درجات أعلى من الشخص (أ) و 12 شخصاً حصلوا على درجات أقل منه، فإننا في هذه الحالة نضع هذا الشخص (أ)

لو أي أشخاص آخرين في مرتبة قريبة منه بين المجموعتين معتبرين أن 27 حالة تقع أعلى منه و 12 حالة تقع أدنى منه. وحيث أن 12 هي 321% من العدد الكلي البالغ 40 شخصاً فلذا في هذه الحالة نقول أن الدرجة المئينية للشخص (أ) هي 31.

وبطريقة الحساب هذه فإن الشخص الذي يقع في منتصف الجماعة تماماً هو الذي يحصل على المئيني الخمسين، وهو ما يعبر عنه إحصائياً باسم الوسيط وهو في الواقع يدل على أداء الفرد للنموذجي في الجماعة. وبغض النظر يمكن الحديث عن مقاييس الترتيب النسبي المماثلة مثل الربيعيات أو العشيريات وغيرها.

ومما لا شك فيه أن تحويل الدرجات الخام إلى درجات مئينية يغير من شكل التوزيع. فمن المعروف أن التوزيع التكراري للنموذجي للدرجات الخام هو المنحني الاعتدالي أو ما هو أقرب إليه. حيث يكون أكبر عدد من الحالات في المنتصف (أو حول المتوسط الحسابي) ويقل عدد الحالات تدريجياً كلما اتجهنا نحو طرفي التوزيع. أما إذا تحولت جميع الدرجات الخام إلى درجات مئينية فإن للتوزيع التكراري يصبح أقرب إلى المنحني المستطيل. ومعنى ذلك أن الحالات الكثيرة التي تقع في المنتصف تنتشر في حين أن الحالات القليلة التي تقع في الأطراف تتجمع بعضها مع بعض.

ولا بد أن نشير إلى أننا لا نستطيع استخراج متوسطات الدرجات المئينية، لأن هذا المتوسط يختلف في مقداره عما لو استخرجنا المتوسط الحسابي للدرجات الخام ثم حولناه إلى درجة مئينية. وبدلاً من اللجوء إلى المتوسط الحسابي لقياس النزعة المركزية للدرجات المئينية يمكن للباحث أن

يعتمد على الوسيط باعتباره المقياس الإحصائي المناسب للنزعة المركزية في هذه الحالة.

كما يجب الإشارة إلى أن الدرجات المتينة لا يمكن المقارنة بينها ما لم تكن الجماعات التي تشتق منها يمكن المقارنة بينها بالفعل. فإذا اختلفت الجماعات، كأن تكون إحداها مثلاً تلاميذ مدرسة ثانوية صناعية والثانية عمال تلمذة صناعية ولثلاثة طلاب كلية الهندسة، فقد تكون الدرجة الخام التي تقابل الدرجة المتينة إلى 60 مثلاً لطلاب الجامعة تقابل مثلاً المتيني 75 في مجموعة طلاب المدرسة لثانوية الصناعية وتقابل كذلك المتيني 90 في مجموعة عمال التلمذة الصناعية. بعبارة أخرى إذا لجأنا إلى استخدام الدرجات المتينة لا بد للباحث أن يضع في اعتباره المجموعة المعيارية أو عينة التلقين التي اشتقت منها هذه المعايير، ولو أن ذلك لا يحول دون استخدام مجموعة من جداول التحويل إلى درجات متينة تسمح للمختبر أو الباحث أن يقارن للشخص الذي يختبره بمجموعات مرجعية مختلفة.

العمر العقلي ونسبة الذكاء:

تعتمد بعض الاختبارات على مقياس للعمر في تفسير درجاتها، ومن أشهر هذه الاختبارات ستانفورد بينيه، وفي هذه الحالة يتم تجميع أسئلة الاختبار حسب مستويات العمر، فمثلاً الأسئلة التي يجيب عليها أغلبية الأطفال من سن 7 سنوات، وتلك التي يجيب عليها أغلبية الأطفال من سن 8 سنوات تعتبر ملائمة لمستوى 8 سنوات وهكذا، ومعنى ذلك أن درجة الطفل في مثال هذا الاختبار تقابل أعلى مستوى عمري يمكنه أن يصل إليه بنجاح. فإذا استطاع طفل عمره 9 سنوات أن يجيب بنجاح على أسئلة عمر 12 سنة يكون عمره للعقلي 12

بصرف للنظر عن عمره الزمني. وهو في هذه الحالة يكون سابقاً لعمره الزمني بثلاث سنوات ما دام أدائه يكافئ أداء طفل متوسط من عمر زمني أكبر منه بثلاث سنوات.

وقد لوحظ أن أداء الفرد في الاختبارات التي تعتمد على مفهوم العمر يظهر مقدراً معيناً من الثبات. بعبارة أخرى قد يفضل الطفل في أداء بعض الأسئلة التي تقع في مستوى أدنى من عمره العقلي، كما يمكنه الإجابة على بعض الأسئلة التي تقع في مستوى أعلى منه. ولذلك نلجأ عادة إلى حساب العمر القاعدي أو الأساسي أي أعلى مستوى عمري يستطيع الطفل أن يجيب على جميع أسئلته والأسئلة السابقة عليه ثم تضاف إليه درجات جزئية، بالطبع في صورة كسور من السنة أو الشهر، تدل على جميع الأسئلة التي أجاب عليها من مستويات أعلى من العمر القاعدي، وبذلك يصبح العمر العقلي للطفل هو مجموع العمر القاعدي والأشهر الإضافية التي تدل على ما اكتسبه الطفل من مستويات عمرية أعلى.

وقد نلجأ إلى استخدام معايير العمل العقلي في اختبارات لا تنقسم إلى مستويات حسب السنوات. وفي هذه الحالة نحصل أولاً على الدرجة الخام التي قد تكون العدد الكلي للأسئلة الصحيحة في الاختبار ككل أو قد تكون معتمدة على الزمن أو عدد الأخطاء أو مزيج منها. ثم نحسب متوسطات للدرجات الخام التي يحصل عليها الأطفال في كل مجموعة عمرية من بين عينة التقيين، وتكون تلك المتوسطات هي معايير العمر لذلك الاختبار.

ويجب أن نؤكد أن وحدة العمر العقلي لا تبقى ثابتة بتقدم العمر ولكنها تميل إلى الضمور في السنوات المتأخرة عنها في السنوات المبكرة من النمو.

فإذا كان الطفل متخلفاً عاماً واحداً في سن 4 سنوات فإنه يصبح متخلفاً ثلاث سنوات في سن 12 سنة. فعلم واحد من " للنمو " العقلي في الفترة من سن 3 حتى 4، يعادل 3 سنوات من النمو في الفترة من سن 9 حتى سن 12، وحيث أن للنمو العقلي يتقدم بسرعة في السنوات المبكرة من النمو ثم يتناقص تدريجياً بالتقارب للفرد من حد النضج، فإن العمر العقلي يضمّر تبعاً للعمر. وهذه العلاقة يمكن فهمها بصورة أوضح إذا تصورنا طول الشخص على أنه يمكن صياغته في صورة " عمر طولي " نجد أن الفرق بالسنتيمترات بين " العمر الطولي " في عمر 3 أو 4 يكون أكبر من ذلك الفرق بين " العمر الطولي " في سن 10 و 11. ومعنى ذلك أن عاماً واحداً من التفوق أو التخلف في سن 5 سنوات يمثل انحرافاً عن المستوى المتوسط أكبر كثيراً من مقدار مماثل له من التفوق أو التخلف في سن 10 سنوات.

ومن أجل الحصول على مقياس يسمح بتفسير موحد بصرف النظر عن العمر الزمني للمفحوص ابتكر ترمان مفهوم نسبة الذكاء رغم أن الحاجة إلى مثل هذه النسبة أشار إليها قبله شترن وكولمان، كما سنوضح فيما بعد، ونسبة الذكاء هي نسبة العمر العقلي إلى العمر الزمني مضروبة في 100 للتخفيف من الكسور العشرية، على النحو الآتي:

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلي}}{\text{العمر الزمني}} \times 100$$

فإذا تعادل العمر العقلي والعمر الزمني تكون نسبة الذكاء في هذه الحالة 100 وتدل على أداء متوسط أو عادي، وإذا قلت عن 100 دلت على تخلف أو

ضعف، وإذا زالت عن 100 دلت على التفوق. وتفيد نسبة الذكاء أيضاً في ضبط الضمور في وحدة العمر العقلي ضبطاً تلقائياً باستخدام للنسبة. فمثلاً إذا كان العمر العقلي لطفل عمره الزماني 4 سنوات هو 3 تصبح نسبة ذكائه في هذه الحالة 75 فإن هذا الطفل حينما يصل إلى سن 12 سنة يصبح عمره العقلي 9 وتظل نسبة ذكائه ثابتة أي 75. أي أن نسبة الذكاء تدل على نفس الوضع النسبي للفرد في الجماعة سواء في سن 4 سنوات أو في سن 12 سنة. وبعبارة أخرى فإن نسبة الذكاء يمكن استعمالها في المقارنة في مختلف الأعمار، بمعنى أن تفسير نسبة الذكاء يظل هو نفسه بصرف النظر عن العمر الزمني للمفحوص.

ولا بد أن نشير في هذا الصدد إلى أن نسبة الذكاء تظل ثابتة في حالة واحدة فقط وهي أن وحدة العمر العقلي تتناقص تناقصاً يتناسب تناسباً مباشراً مع العمر. وهذا شرط ضروري لتمكين استخدام نسب الذكاء استخداماً له معنى. ومعنى ذلك أنه حينما تتناقص وحدة العمر العقلي فإن الفروق الفردية التي تقاس بهذه الوحدة سوف تتزايد بنفس النسبة. ويشبه ذلك من الناحية الحسابية أن الفروق الفردية في الطول تكون أكبر 12 مرة حينما تقاس بالبوصات عنها لو قيست بالأقدام وذلك لأن البوصة هي $\frac{1}{13}$ من القدم وب نفس المطلق فإن الفروق الفردية في العمر العقلي يجب أن تكون في سن 14 ضعفها سن 7 سنوات، ولا يمكن لنسب الذكاء أن تظل ثابتة إلا إذا تزايد الانحراف المعياري للأعمال العقلية بنفس نسبة زيادة الأعمار الزمنية، وبهذا الشرط وحده يمكن أن يكون لنسبة الذكاء نفس المعنى في مختلف الأعمار.

وقد أمكن تحقيق هذا الشرط في اختبار ستانفورد بينيه فقد اتضح من دراسة انتشار الأعمار العقلية من سن 6 حتى سن 18 أن هناك اتجاهًا نحو زيادة للتباين في الأعمار العقلية تبعاً لزيادة الأعمار الزمنية، ولكننا نجد في كثير من الأحيان عدم تساوي الانحرافات المعيارية لنسب الذكاء في مختلف الأعمار، وهذه الصعوبة لمكن للتغلب عليها في اختبار ستانفورد بينيه بإعداد جدول تصحيح يستخدم فيه تصحيح نسب الذكاء في مستويات عمرية معينة.

وهناك مشكلة أخرى في تطبيق معيار نسبة الذكاء نجدها في اختبار البالغين والكبار. فالواقع أن مفهوم معايير العمر يدل دلالة مباشرة على أن فائقته مقصورة على الأطفال. أما بالنسبة للكبار فإننا نجد أن في اختبار مثل ستانفورد بينيه لا يتحسن أداء البالغ المتوسط كثيراً أبعد من مستوى 15 سنة. وعلى ذلك فإن العمر العقلي للبالغ المتوسط هو حول هذا المقدار (حوالي 15 سنة و 9 أشهر). وقد أضيف إلى الاختبار مستويات رابعة لذكاء الكبار حتى يتوفر في الاختبار سقف مناسب ويلتقي يمكننا الحصول على أعمار عقلية أعلى من 15 سنة و 9 أشهر، ولكننا إذا وجدنا شخصاً معيناً يحصل على عمر عقلي مقداره 20 في هذا الاختبار فإن ذلك لا يسمح لنا بتفسير مؤكد وعلى نفس الدرجة من الوضع كما نجد في حالة الأعمار العقلية 8 أو 10. فمن المؤكد أن العمر العقلي 20 لا يمكن تحديده بما يمكن أن يؤديه الشخص المتوسط للبالغ من العمر 20 سنة، لأن الشخص المتوسط للبالغ من العمر 20 سنة يحصل على عمر عقلي مقداره 15 سنة و 9 أشهر. أما إذا اختبارنا شخصاً بالغاً ضعيف العقل يقل عمره العقلي عن 15 سنة و 9 أشهر فإن مفهوم العمر العقلي لا يختلف في استخدامه في هذه الحالة عنه حينما نستخدمه مع الأطفال. أما في حالة البالغين الأسوياء أو المتفوقين فلا مناص من استخدام أنواع أخرى من الدرجات

مثل المئينيات (التي أشرنا إليها) أو الدرجات المعيارية (التي سنشير إليها فيما بعد).

أما المشكلة الثالثة التي تولجها حينما نفسر الدرجات الخام بدرجات العمر تتمثل في أن هذا النوع من التعبير لا يستخدم إلا في الوظائف النفسية والعمليات العقلية التي تتغير تغيراً واضحاً ومتسقاً مع العمر والنمو. أما بالنسبة للمسات التي لها علاقة ضئيلة أو لتي لا تكون لها علاقة بالعمر للزمني لا يجوز مطلقاً تحويل الدرجات الخام لمقاييسها إلى درجات أعمار. ومن أمثلة الاختبارات التي يستحيل تطبيق معايير الأعمار عليها الاختبارات التي تقيس سمات الشخصية وخصائصها.

للدرجات المعيارية:

تعتبر الدرجات المعيارية أفضل صورة تحويل للدرجات الخام. وهي تعبر عن بعد الدرجة التي يحصل عليها الفرد عن المتوسط في ضوء الانحراف المعياري للتوزيع. وللحصول على الدرجة المعيارية لا بد أولاً من حساب المتوسط الحسابي الذي هو للوسط الحسابي الذي يحصل عليه جميع الدرجات ثم قسمة هذا المجموع على عدد الأفراد. ثم حساب الانحراف المعياري (الذي هو مقياس للتشتت أو الانتشار) وهو في أصله متوسط الانحرافات درجات المفحوصين كأفراد عن المتوسط الحسابي للجماعة. ويدل الانحراف المعياري على مقدار الاختلاف داخل الجماعة، على عكس المتوسط الذي يدل على للزعة المركزية للدرجات. وبعد حساب كل من المتوسط والانحراف المعياري يمكن إيجاد الفرق بين الدرجة التي يحصل عليها الفرد ومتوسط الجماعة ثم قسمة هذا الفرق على الانحراف المعياري للمجموعة، وبهذه الطريقة

نستبدل الدرجة الخام بدرجة مشتقة نزل على وضع الفرد بالنسبة لمتوسط الجماعة، وبفقدنا الانحراف المعياري في تحديد مقدار البعد عن هذا المتوسط ارتفاعاً أو انخفاضاً، وبالتالي نستطيع أن نقول أن الفرد (أ) أعلى من المتوسط بمقدار 205 وحدة لانحراف معياري، وأن الفرد (ب) أقل من المتوسط بمقدار 1- وحدة لانحراف معياري.

ومن المشكلات التي تعانيها الدرجات المعيارية أن بعض قيمها قد تكون سالبة كما أنها تتضمن كسوراً من الواحد الصحيح مما يسبب خلطاً وصعوبة في استخدامها سواء في الجانب الحسابي الخاص بحساب هذه الدرجات أو الجانب الخاص بتفسير هذه الدرجات والاستفادة بها عملياً. ولهذا السبب اقترحت عدة تعديلات من أجل وضع الدرجات المعيارية في صورة أكثر ملاءمة، منها مثلاً الدرجات الثنائية التي تفترض متوسطاً مقداره 50 وأن كل وحدة انحراف معياري تساوي 10. (ومن المعروف أن متوسط الدرجات المعيارية صفر وأن انحرافها المعياري هي واحد صحيح).

اختبارات الذكاء العام

اختبارات الذكاء العام

• توجد فئة من الاختبارات النفسية تسمى اختبارات الذكاء العام. وتتميز هذه الاختبارات بأنها تستخدم في مواقف متنوعة ويتحدد صدقها في ضوء محكات أكثر شمولاً. كما تتميز هذه الاختبارات بأنها تعطي درجة واحدة مثل نسبة الذكاء، لتدل على المستوى العقلي العام للمفحوص، ويرى ترمان وميريل أن الوصول إلى هذا التقدير الشمولي للمستوى العقلي يتطلب " تصويب الرماح إلى الأهداف الحاسمة ". وبعبارة أخرى يعرض على المفحوص عدداً كبيراً ومتنوعاً من الأعمال على أساس أنها تمثل عينة ملائمة من أساليب الأداء العقلي. ومع ذلك فإن للبحوث الحديثة تؤكد أن هذه الاختبارات تتشعب بوظائف عقلية معينة مثل القدرة اللغوية، وتتجاهل تماماً بعض الوظائف العقلية الأخرى.

• وكثيراً ما تسمى اختبارات الذكاء العام باختبارات الاستعداد للمدرسي لأن صدقها يتحدد في العادة في ضوء محكات التحصيل الأكاديمي. كما أن هذه الاختبارات تستخدم في أغلب الأحوال في أغراض للتصنيف للمبتدئة العامة (سواء من الناحية التربوية أو المهنية أو العسكرية)، والمفروض أن يتبعها تطبيق اختبارات الاستعدادات الأكثر تخصصاً. ويشيع هذا على وجه الخصوص في اختبار المراهقين والراشدين الأسوياء عند الإرشاد أو الانتقاء أو للتصنيف أو للتوجيه أو غير ذلك من الأغراض التي أشرنا إليها في الفصل الأول من هذا الكتاب، كما يوجد استخدام آخر لاختبارات الذكاء في الأغراض الإكلينيكية خاصة وهو تحديد النصف العقلي وتشخيصه. وفي مثل هذه الأحوال فإن استخدام الاختبارات الفردية مثل مقولس ستانفورد بينيه ومقواس وكسملر أكثر شيوعاً من غيره.

وسوف نتناول في هذه الوحدة المقاييسين بشيء من التفصيل، دون أن نتجاهل الاختبارات الأخرى التي سعت إلى قياس الذكاء العام.

الفرد بينيه

الفرد بينيه هو العالم لغذ الأصل، بدأ تدريجه في ميدان الطب ثم سرعان ما أصبح أشهر علماء النفس في فرنسا في عصره، وله فضله للعظيم في علم النفس في بلده. فقد أنشأ زميله بيوني أول معمل لعلم النفس عام 1889 في جامعة السوربون، وأول مجلة متخصصة باللغة الفرنسية "حولية علم النفس عام 1895".

ولقد بدأ بينيه عالماً تجريبياً، فقبل أن يعد اختباره العقلية للمشهوره اهتم بدراسة العمليات العقلية المعرفية، وأطلق على منهجه اسم " سيكولوجية الفرد " وفي رأيه أن أهم ما نجده من فوارق بين الناس إنما هو في الأنشطة العقلية العليا كالاستدلال والحكم، وهي جميعاً من المسائل التي يصعب قياسها بالوحدات الفيزيائية (كالإحساس) كما لا تتوافر فيها الوحدات الأولية للسلطان الذهني، والتي كان يهتم بها أغلب علماء النفس من أنصار بنائية فوننت (وملهم تتشورا).

وقد تأكدت صحة وجهة نظر بينيه من نتائج بعض الدراسات التي أجريت على صدق اختبارات جيمس ماكين كاتل (التي لفتت فيها أنثر السيد فرانسيس جالتون في إنجلترا) كمقاييس عقلية، وبخاصة بحوث كلارك وزلو فرانسيس جالتون في إنجلترا). ولذلك بدأ بينيه يطبق في بحوثه عن الاختبارات العقلية ما توصل إليه من نتائج البحث للتجريبي في عملية التفكير، فانتقد

لاختبارات كاتل وجالتون لأنها تقتصر على النواحي الحسية البسيطة، وفضل الاختبارات الأكثر تعقداً وتركيباً، واقترح عدداً من الاختبارات لقياس " الملكات العقلية " .

وقضى بينيه للفترة بين عامي 1896 و 1905 في إجراء سلسلة طويلة من البحوث على الأطفال الأسوياء وغيرهم مجرباً أنواعاً عديدة من الاختبارات لمختلف " الملكات " ، بالإضافة إلى عدد من التجارب حول المسائل الأخلاقية في سيكولوجية التفكير في ذلك الوقت مثل التجريد، والتفكير بدون صور وزمن المرجع والتمييز الحسي. وكان الهدف الذي يسعى إليه بينيه من ذلك هو تحديد العلاقة بين هذه الاختبارات والمصر للزملي والتحصيل المدرسي وتقديرات المعلمين للفترة العقلية. وكانت بحوثه جيدة للتصميم ولأت به إلى التحول من الاهتمام بالاختبارات التي تقيس " الملكات " التقليدية إلى تلك التي تقيس المستوى العقلي بالفعل.

وهكذا كانت نظرة بينيه إلى الذكاء أكثر شمولاً، حين لجأ إلى استخدام درجة كلية واحدة (المصر العقلي) في قياس الذكاء.

نشأة مقياس بينيه:

في عام 1904 طلبت وزارة المعارف الفرنسية من لفرد بينيه والطبيب لفرنسي ثيودور سيمون دراسة الطرق التي يمكن أن تستخدم في تصنيف الأطفال الذين يتميزون ببطء التعلم في المدارس الفرنسية، فعكفا على وضع مقياسهما للمشهور للذكاء بغرض للتمييز بين الأطفال الأسوياء وضعفاء العقل.

ولم يكن الهدف دراسة وتحليل استعدادات المتخلفين عقلياً، وإنما كان الهدف عملياً بحثاً وهو تقويم وقياس ذكاء هؤلاء الأطفال بصفة عامة وتحديد مستواهم العقلي ومقارنتهم بالأطفال الأسوياء من نفس العمر والمستوى. أي الهدف كما يقول بينيه هو بناء مقياس " مقري " للذكاء يتكون من عدد من (الاختبارات) المنتزجة في الصعوبة، مرتبة حسب الأعمار حتى يستطيع الفاحص أن يحدد مدى تقدم الطفل عقلياً أو تخلفه. وقد سعى بينيه إلى أن تكون هذه " الاختبارات " متقوعة، مع التركيز على فئات معينة من العمليات العقلية اهتم بها دون غيرها عدد إعداد المقياس وهي: الحكم والفهم للعام والمبادأة والقدرة على التكيف.

ويتكون مقياس بينيه الأصلي الذي نشر عام 1905 بلسن نقياس بيتيغ - ميمون للذكاء من 30 اختباراً، وفي يلي قائمة بهذه الاختبارات، نذكرها كاملة لنبين إلى أي حد اختلفت عن احتمالات علماء القياس العقلي قبل بينيه، وكيف أنها هبطت الموقف لظهور عدد كبير من الاختبارات العقلية تتحدر في معظمها " من صلب " مقياس بينيه على حد تعبير أمتنازي بالإضافة إلى شيوعها في معظم تعديلات مقياس بينيه ذاته.

- 1- لتآزر البصري (تتبع شمعة مضيئة بالראس والعينين).
- 2- الفهم عن طريق اللمس (الإصااك بمكعب بعد لعمه).
- 3- الفهم عن طريق البصر أو الروية.
- 4- التعرف على الطعام (التمييز بين قطعة للخشب وقطعة الشكولاته).
- 5- لبحث عن الطعام (الاستجابة لقطعة من الشوكولاته ملفوفة في الورق).
- 6- تتبع التعليمات البسيطة أو الإيماءات المتكررة.

- 7- الإشارة إلى الأياء وخالصة أعضاء الجسم (الرأس، الألف ... إلخ).
- 8- التعرف على الأشياء من صورها.
- 9- تسمية الأشياء في الصورة.
- 10- للتمييز الحسي بين طول خطين.
- 11- تكرار ثلاثة أرقام.
- 12- للتمييز بين وزن شيئين.
- 13- القابلية للإحياء (المقارنة بين خطين متساويين، أو البحث عن أشياء لا وجود لها في صورة).
- 14- تعريف الكلمات البسيطة (التي تدل على أشياء مألوفة).
- 15- تكرار جملة تتكون من 15 كلمة.
- 16- بيان أوجه الاختلاف بين شيئين.
- 17- للذاكرة البصرية.
- 18- رسم الأشكال من الذاكرة.
- 19- مدى ذاكرة الأرقام.
- 20- بيان أوجه التشابه بين الأشياء.
- 21- التمييز بين الخطوط بسرعة.
- 22- ترتيب 5 أوزان.
- 23- تحديد الوزن الناقص (من بين الأوزان الخمس في الاختبار السابق).
- 24- إعطاء كلمات ذات قافية واحدة.
- 25- تكلمة الجمل.
- 26- بناء جملة تحتوي على 3 كلمات معينة.
- 27- الإجابة عن الأسئلة (مثل: ماذا تفعل وأنت نائم؟).
- 28- تحديد الوقت بعد أن نحل عقارب الساعة محل بعضها.

29- شي وتطبيع قطعة من الورق.

30- للتمييز بين الكلمات المجردة (مثل: حزن وسأم).

وإذا فحصنا هذه القائمة فحصاً دقيقاً نجد أن الاختبارات للثلاث الأولى هي في الواقع اختبارات للنمو الحركي، أما الاختبارات الأخرى فيمكن أن تعد من نوع اختبارات الأداء الأكصي. ومن بين هذه الاختبارات نجد 18 اختباراً (حوالي الثلثين) تقيس ما يسميه جيلفورد التفكير المعرفي، كما نجده اختبارات تقيس الذاكرة بأنماطها المختلفة. ولا يوجد إلا 3 اختبارات فقط تقيس ما يسمى في الوقت الحاضر قدرات التفكير الابتكاري (أو الإنتاج للتباعدي عند جيلفورد).

وكانت الطريقة التي استخدمها بينيه في تحليل مقايسه غير ملائمة، فقد طبقه على 50 طفلاً (10 من كل فئة من فئات الأعمار 3، 5، 7، 9، 11) اختارهم مملوهم على أساس أنهم متوسطو القدرة العقلية، بالإضافة إلى عدد لم يذكره بينيه من الأطفال ضعاف العقول. وكانت طريقة التصحيح تعتمد على أساس تجريبي مبني يسمح بتصنيف الأطفال تبعاً لأعلى مستوى يصل إليه الطفل، فالمعتوهون لا يمكنهم تجاوز الاختبار رقم 6. ولا يتجاوز للبلهاء الاختبار رقم 15، أما حدود المألوتين فلم تكن معروفة، رغم أنهم يعجزون في العادة عن الإجابة على الأسئلة المجردة التي يستطيع معظم الأطفال للعاليين من سن 11 الإجابة عليها.

من المحتمل أن ننجح في يوم ما في إيجاد علامات على التخلف الميكولوجي مستقلة تماماً عن العمر. وقد يكون من المفيد جداً التعرف على هذه العلامات إلا أننا في الوقت الحاضر نجد أن ما يجذب انتباهنا أكثر من غيره هو التشابه بين الأسوياء من صفار السن وغير الأسوياء من كبار السن. وهذا

لانتسابه من الكثرة والغلبة بحيث أننا لو حصلنا على استجابات طفل لا نعرف عمره لا نستطيع الحكم عليه بالسواء أو عدم السواء^١.

التحليل السيكولوجي للتقليدي مثل الملكات أو عتبات الإحساس أو غيرها. وبالإضافة إلى ذلك فإن وضع اختبار في المقياس أو حذفه أو تعديله كان من المسائل التجريبية، ويدل على خبرة بينيه الطويلة بالتجريب العملي. فقد كان يهتم بمدى العلاقة بين الاختبار وأحكام المحك (أي بلوغ من الصدق). ثم إن هذه الاختبارات لم تكن محدلات معملية دقيقة وإنما كانت أدوات عملية بسيطة وهامة لدراسة الأطفال في مواقف واقعية (مثل حجرة للدراسة).

مقياس ستانفورد - بينيه لترمان

لم تكن محاولة جودارد ترجمة ونقل مقياس بينيه إلى اللغة الإنجليزية هي المحاولة الوحيدة، ففي عام 1910 نشر هوى ووربل ترجمة له، كما نشر فالين وكولمان ترجمة أخرى في عام 1911. وفي عام 1913 بدأ بيرت نقل المقياس إلى الإنجليزية وتقنيه على عينة كبيرة من الأطفال الإنجليز عندما عين أخصائياً نفسياً لمقاطعة في لندن، وظل يتابع مشروعه هذا طوال تاريخه العلمي كأستاذ لعلم النفس بجامعة لندن. بالإضافة إلى نقل المقياس إلى إيطاليا وألمانيا وروسيا والسويد واليابان في عام 1920 ترجم أحمد فكري مقياس بينيه الأصلي إلى اللغة العربية. وقدم له تعليمات تقتضية، وكان يعتقد خطأ أن المدرس - دون تدريب سابق - يمكنه تطبيق المقياس والحكم على نكاء الطفل. ومع ذلك فإن أكثر هذه المحاولات شهرة في تاريخ لقياس النفسي محاولة لويس ترمان (1877-1956) أستاذ علم النفس بجامعة ستانفورد، والتي عرف بها المقياس باسم مقياس ستانفورد بينيه للنكاء.

وقد ظهرت الطبعة الأولى لهذا المقياس عام 1916. وأدخل ترمان على المقياس الأصلي عدداً من التعديلات والإضافات بحيث يمكن القول أنه اختبار جديد، فقد كان حوالي ثلث الأسئلة جديداً، وعدل ترمان عدداً من الأسئلة القديمة وأعيد توزيعها على المستويات العمرية المختلفة أو حُلِّفت. وأعيد تقنين المقياس كله على عينة من الأمريكيين بلغ عددها 1000 طفل و400 راشد. وأعدت تعليمات مفصلة لتطبيق كل اختبار وتصحيحه، واستخدم معيار نسبة الذكاء (تعديلاً لمعيار للنسبة العقلية لشترن وكولمان).

وقد كتبت مجلة المقطف المصرية مقالاً عن هذا المقياس عام 1927 كما كتب عنه الدكتور أمير بقطر عام 1928، ونقله إلى اللغة العربية لأول مرة طبيب غير متخصص في علم النفس هو الدكتور حسن عمر.

وفي عام 1937 ظهرت الطبعة الثانية المعدلة للمقياس، وتتكون من صورتين متكاملتين، الصورة (ل) والصورة (م)، وفي هذه الطبعة زاد عدد أسئلة المقياس وأعيد تقنيته تماماً على عينة جديدة جيدة الانتقاء من المجتمع الأمريكي بلغ عددها أكثر من 3000 مفحوص تتراوح أعمارهم بين سن 1 ½ إلى سن 18. واشتملت كل عينة عمرية على عدد متساوٍ من الذكور والإناث. وتم اختبار معظم المفحوصين من سن 6 سنوات وما بعدها من المدارس، أما عينات أطفال ما قبل المدرسة (عينات الأعمار 1 ½، 2، 2.5، 3، 3.5، 4، 4.5، 5، 5.5 سنة) فلم الحصول عليها بطرق مختلفة، أيسرها تطبيق المقياس على أخوة أطفال وتلاميذ المدارس، وقد روعي في انتقاء العينات للتوزيع الجغرافي والمستويات الاقتصادية والاجتماعية.

وفي عام 1960 ظهرت طبعة ثالثة لمقياس ستانفورد بينيه في صورة واحدة فقط تجمع أفضل الأسئلة من صورتني طبعة عام 1937، وبدون تقديم أي محتوى جديد أو صياغة أسئلة جديدتكان من الممكن حذف الأسئلة التي لا تساير العصر، وأبعد توزيع الأسئلة التي تغيرت مستويات صعوبتها في الفترة بين التعديلين تبعاً للتغيرات الثقافية، وفي إعداد طبعة عام 1960 ولجه ترمان مشكلة شائعة في لقياس النفسي وتتلخص في أنه من المستحب عامة إخال تعديلات دورية على المقاييس للإفادة من التقدم العلمي في ميدان بناء الاختبارات، ومن الخبرة السابقة باستخدام الاختبار، بالإضافة إلى جعل محتوى الاختبار مسلياً لروح العصر. وتردد الحاجة إلى الاعتبار الأخير بالنسبة لأسئلة المعلومات والمواد المصورة والتي تتغير تبعاً لتغيير الظروف الثقافية، فاستخدام أسئلة قديمة يؤثر في علاقة الفاحص بالمفحوص وفي الصدق الظاهري للاختبار قد يغير في مستوى صعوبة الأسئلة. ومن ناحية أخرى فإن التعديل قد يجعل معظم البيانات المتوافرة عن الاختبار لا ترتبط بالصورة الجديدة، فالاختبارات التي استخدمت استخدامات واسعة لعدة سنوات كمقياس ستانفورد بينيه تتوافر عنها ثروة من المعلومات يجب أن توزن في مقابل الحاجة إلى التعديل. ولهذه الأسباب فضل تلاميذ ترمان الجمع بين الصورتين في صورة واحدة بعد انتقاء أفضل ما فيهما. وفقدان صورة مكلفة ليس ثمناً فاحشاً لتحقيق هذه الغاية، وخاصة أنه أصبحت توجد في الوقت الحاضر اختبارات فردية أخرى جيدة الصنع لم تكن متوافرة عام 1937.

تحليل اختبارات المقياس

لقد استغرق تعديل عام 1937 لمقياس ستانفورد بينيه أكثر من عشر سنوات من البحث العلمي المنظم حول تجميع الأسئلة الجيدة وتجريبها على جماعات صغيرة العدد من المفحوصين الذين طبق عليهم مقياس عام 1916.

وكانت الخطوة الأساسية في ذلك هي تطبيق الصورتين المبدئيتين في ذلك الوقت على عينة للتقنين (3184 مفحوصاً) ومنها تم الانتقاء النهائي للأسئلة وتوزيعها على الأعمار وقد أعطيت لكل مفحوص صورتاً الاختبار بفواصل زمنية يتراوح بين يوم واحد وأسبوع واحد. وفي هذه الخطوة كان محك تحليل الأسئلة لتحديد صحتها هو العمر الزمني والدرجة الكلية للمركبة من الصورتين المبدئيتين. ويرتبط المحك الثاني بما أسميناه الاتساق الداخلي حتى يتحقق في المقياس درجة من التجانس. واستخدمت ثلاثة مقاييس في تحليل كل سؤال (اختبار) في المقياس هي:

- 1- منحنيات النسب المئوية للمفحوصين الذين يجتازون السؤال بنجاح في الأعمار الزمنية المتتالية.
- 2- منحنيات للسبب المئوية للمفحوصين الذين يجتازون السؤال بنجاح في الفئات المتتالية للدرجة الكلية في المقياس (وبالطبع لكل صورة من صورتين المقياس منحنياتها الخاصة).
- 3- معامل الارتباط التثقي بين كل سؤال (اختبار) والدرجة الكلية في المقياس.

ونتفق للطريقة الأولى مع أسلوب بينيه الذي كان يستخدمه في تحليل المفردات وانتقاء الأسئلة التي تصلح للمستويات العمرية المختلفة. وفي هذا يجب أن نشير إلى أن ترمان استخدم للحكم على صلاحية الأسئلة نمباً مئوية في الأعمار الصغيرة أعلى منها في الأعمار الكبيرة. ومن ذلك مثلاً أنه كان يعتبر السؤال مناسباً لمن 3 سنوات إذا أجاب عليه 73% من أطفال هذا السن إجابة صحيحة، أما في حالة الأطفال من سن 10 سنوات فكانت هذه النسبة 59%.

ويشير ماكنمر إلى أن هذا الاختلاف في النسب المتوقعة هو من المطالب للضرورة في مقاييس العمر. فمن الملاحظ أن الانحراف المعياري للعمر العقلي يجب أن يتزايد مع العمر إذا كان لا بد للنسبة الذكاء أن تظل ثابتة. وعلى ذلك فإذا كان التباين في العمر العقلي في الأعمار العليا أكبر منه في الأعمار الدنيا فإن ذلك يعطي ضرورة وجود انتشار أكبر أداء المفحوصين للكل في مستويات الأعمار المتجاورة. وعلى ذلك فمع الزيادة في العمر يقل عدد المفحوصين الذين يجتازون بنجاح. السؤال المطابق لمستواهم العمري.

ويمكن أن نوضح هذه المسألة بمثال بسيط لنفرض أنه بين الأطفال من سن 5 سنوات يصل 40% منهم إلى العمر العقلي 5، بينما يصل 30% إلى أعمال أعلى، و 30% آخرون إلى أعمار أقل من هذا المستوى. وحيث أن أي مفحوص عمره العقلي 5 سنوات أو أكبر يجب أن يجتاز من الوجهة النظرية السؤال الخاص بسن 5 سنوات فإن 70% (40% + 30%) من الأطفال من فئة 5 سنوات نتوقع نجاحهم في أسئلة سن 5 سنوات. ولنفر من ناحية أخرى أنه من بين الأطفال من سن 10 سنوات لا يصل إلى العمر العقلي 10 سنوات إلا 20%، بينما يصل إلى أعمار أعلى من ذلك 40% منهم، وإلى أعمار أقل من ذلك 40% آخرون. وبالمطبع لا يمكن الحصول على مثل هذا التوزيع إلا إذا كان التباين في العمر العقلي للأطفال سن 10 سنوات أكبر منه في سن 5 سنوات. ومعنى ذلك أنه من بين الأطفال من سن 10 سنوات تصبح النسبة المتوقعة لاجتياز الأسئلة المخصصة لهذا السن هي (20% + 40% = 100%). وفي هذا المثال تصبح النسبة المتوقعة التي يمكن استخدامها محكاً لانتقاء الأسئلة في سن 5 سنوات هي 70%، وفي سن 10 سنوات هي 60%.

وفي الممارسة الواقعية لانتقاء اختبارات مقياس ستانفورد بينيه نجد أن المحك يتناقص من 77% عند مستوى من عامين إلى ما هو أقل قليلاً من 50% في مستوى الرشد العادي. وتقل للنسب عن هذا في انتقاء أسئلة مستويات الرشد المتفوق وذلك للوصول إلى سقف مناسب للاختبار. وعند الانتقاء النهائي للمفردات لم يقتصر الاهتمام على تمايز العمر والاتساق الداخلي لحسب، وإنما امتد ليشمل أيضاً مدى تناقص الفروق بين الجنسين وتوازنها في للنسب المئوية للناجحين في مختلف المستويات العمرية . واستبعدت بالفعل المفردات التي اختلف فيها أداء أحد الجنسين عن الآخر بنسب لها دلالة إحصائية، على أساس أن مثل هذه المفردات قد تدل على تباين اصطناعي أو غير مرتبط بما يقيسه الاختبار نتيجة للخبرات السابقة للجنسين.

وفي إعداد طبعة عام 1960 من المقياس لاختبرت الأسئلة من الصورتين ل، م على أساس أداء 4498 مفحوصاً تتراوح أعمارهم بين عامين، 18 سنة الذين أجابوا على صورتَي المقياس أو إحداهما بين عامي 1950، 1954. كما تضمنت العينة مجموعتين طبقيتين اشتملتا على 100 طفل من سن 6 سنوات، و100 مراهق من سن 15 سنة، صنفتا حسب مهنة الوالد ومستواه التعليمي.

وبعني هذا أن طبعة 1960 لم تتضمن إعادة لتقنين المقياس. واستخدمت العينات الجديدة لمراجعة للتغيرات في صعوبة الأسئلة فقط بعد ما تحددت مستويات الصعوبة بإيجاد للنسب المئوية للأطفال الذين يجتازون السؤال في أعمار عقلية متتابعة في صورتَي عام 1937. كما روجعت بعض الأسئلة للتحقق من احتمال وجود فروق جغرافية أو اقتصادية اجتماعية وذلك من خلال

المقارنة بين المجموعات الفرعية التي تكون العينة الكلية. ولم يتضمن المقياس كما أُرنا، أي مواد جديدة، وإنما عدلت بعض الرسوم الخاصة بالأشياء الشائعة ممايرة للعصر. وفي هذه الأحوال اختبرت الاختبارات المعدلة قبلاً على عينات صغيرة قبل تضمينها في المقياس.

وصف المقياس

يتألف مقياس متلفرد ببنية من صندوق مجموعة من اللعب المقننة تستخدم مع الأعمار للدنيا، وكثييين من البطاقات المطبوعة، وكراسة لتسجيل الإجابات، وكراسة التعليمات، وكراسة معليير للتصحيح.

وقد رُتبت أسئلة (اختبارات) للمقياس وجمعت تبعاً لمستويات الأعمار التي تمتد من سن عامين حتى سن للرشد المتفوق. وتتحدد مستويات للمقياس من سن 2 حتى سن 5 سنوات في فئات نصف سنوية على النحو الآتي:

2، 6-2، 3، 4-3، 6-4، 5 (مع ملاحظة أن الرمز 6-2، 6-3، 4-6، 4-6، يدل على سن عامين و6 شهور، 3 سنوات و6 شهور، 4 سنوات و6 شهور على التوالي) وذلك بسبب التقدم العقلي السريع في الأعمار للدنيا. أما الفترات من سن 5 سنوات حتى 14 سنة فتتحدد مستويات الأعمار فيها في فئات سنوية، وتكثل للمستويات الأخرى في المقياس على الراشد المتوسط، والراشد للمتفوق. وكان عدد أسئلة (اختبارات) كل مستوى عمري 6 اختبارات، باستثناء مستوى الراشد المتوسط الذي يبلغ عدد أسئلته 8 اختبارات.

وكانت الأسئلة في كل مستوى عمري من مستوى صعوبة موحدة تقريباً، وترتب بصرف للنظر عن أي فروق طفيفة في الصعوبة. وقد أعد سؤال

احتياطي في كل مستوى عمري معادل في صعوبته لأسئلة العمر. ويستخدم هذا السؤال الاحتياطي إذا تطلب الأمر استبعاد أحد الأسئلة العادية بسبب بعض الظروف الخاصة تجعله غير ملائم للمفحوص، أو بسبب عدم الالتزام في التطبيق بحرفية التعليمات بما يتعارض مع تقنين المقياس. وهكذا يبلغ العدد الكلي لاختبارات المقياس 142 اختباراً (122 اختباراً أصلياً، 20 اختباراً احتياطياً).

وقد اختير 4 أسئلة من كل مستوى عمري على أساس الصدق والتمثيل ليتألف منها المقياس المختصر الذي يستخدم حين لا يسمح الوقت بإعطاء المقياس الكلي، وتؤكد المقارنات بين نسب الذكاء كما تتحدد بالمقياس الكلي والمقياس المختصر تطابقاً كبيراً بينهما، بحيث يصل معامل الارتباط بينهما إلى مستوى معامل ثبات المقياس الكلي.

تطبيق المقياس

يتطلب هنا المقياس، كغيره من اختبارات الذكاء الفردية، فاحصاً على درجة كبيرة من المهارة والتدريب لأن إعطاء اختباره وتصحيحها على درجة كبيرة من الدقة ويستلزم أن تتوفر في الفاحص ألفة شديدة بها وخبرة طويلة في التطبيق، لأن تردد الفاحص أثناء التطبيق قد يؤدي إلى فقدان العلاقة الطيبة بينه وبين المفحوص، بل وفقدان ثقة المفحوص به. بل إن أدنى تغيير في صياغة الأسئلة قد يؤدي إلى تغيير في مستوى صعوبتها كما تتحدد في عينة للتقنين. بالإضافة إلى ما يتطلبه المقياس من تصحيح فوري لاختباره عقب إعطائها مباشرة، لأن تتابع إعطاء الاختبارات يعتمد على أداء المفحوص في المستويات السابقة.

ويرى كثير من علماء النفس الإكلينيكيين أن مقياس ستانفورد بينيه ليس مقياساً مقنناً للكفاءة فحسب، وإنما هو كذلك نوع من المقابلة الإكلينيكية، لأنه يهيئ فرصة للتفاعل بين الفاحص والمفحوص، ويهيئ مصادر أخرى للدلالات الإكلينيكية لدى الأخصائي الإكلينيكي المتمرس. بل إن مقياس ستانفورد بينيه، أكثر من غيره من الاختبارات الفردية، يهيئ الفرصة للفاحص أن يلاحظ للطرق التي يلجأ إليها المفحوص في العمل، وأساليب حله للمشكلات، وغير ذلك من الأساليب الكيفية للأداء، بل إن الفاحص يمكنه أن يحكم من خلال الموقف الاختباري الذي يستخدم فيه هذا المقياس على بعض خصائص الشخصية مثل مستوى النشاط والثقة بالنفس والمثابرة والقدرة على التركيز. وبالطبع فإن أي ملاحظات كيفية يمكن تسجيلها أثناء موقف الاختبار يجب تحديدها في ذاتها، ولا يجب تفسيرها على نفس النحو الذي تفسر به درجات الاختبار الموضوعي. وتعتمد قيمة هذه الملاحظات الكيفية إلى حد كبير على مهارة الفاحص وخبرته وإعداداته السيكولوجي العميق، بالإضافة إلى وعيه الشديد بمزالق وحدود هذا النوع من الملاحظات المباشرة. وقد استطاع مورياتري أن يسجل أنماط الملاحظات الإكلينيكية في موقف الاختبار الفردي، على أساس أن جلسة الاختبار هي فرصة لدراسة سلوك الطفل أثناء مواجهة موقف تسود فيه عوامل التحدي والاستثارة والصعوبة والإحباط.

وفي إعطاء مقياس ستانفورد بينيه لا يحل محل المفحوص جميع أسئلة المقياس، وإنما يختبر المفحوص فقط حول مدى المستويات العمرية للملائمة لمستواه العقلي. ولا يتطلب المقياس أكثر من 30-40 دقيقة للأطفال للصغار، ولا يتطلبه أكثر من ساعة ونصف للراشدين. والإجراء المقنن هو أن نبدا الاختبار من مستوى أدنى قليلاً من العمر العقلي المتوقع للمفحوص وعلى ذلك

فإن الاختبارات الأولى يجب أن تكون سهلة بالقدر الذي يستثير في المفحوص الثقة بالنفس، بشرط ألا تكون معسلة جداً بحيث تؤدي بالمفحوص إلى الاستهانة بالاختبار أو الملل. فإذا فشل المفحوص في أي اختبار من اختبارات المستوى الذي يبدأ به يعطى اختبارات المستوى الأدنى منه، وهكذا يستمر الفاحص حتى يصل إلى المستوى الذي يستطيع المفحوص أن يجيب على جميع أسئلته، ويصبح هذا المستوى هو " العمر القاعدي "، ثم يعطيه الاختبارات العليا حتى يصل إلى المستوى الذي يفشل المفحوص في الإجابة على جميع أسئلته ويسمى في هذه الحالة " الحد الأعلى للعمر ". وحين يصل الفاحص إلى هذا المستوى يتوقف عن إعطاء المقياس.

تصحيح المقياس

تصحح جميع أسئلة (اختبارات) مقياس ستانفورد بيليه على أساس النجاح أو الفشل (واحد أو صفر). وتحدد كراسة التعليمات الأداء المطلوب " النجاح " في كل سؤال. وقد تظهر نفس الأسئلة في مستويات عمرية مختلفة ولكنها تصحح بمستويات مختلفة للنجاح وفي هذه الحالة تطبق الاختبارات مرة واحدة ثم يحدد أداء المفحوص المستوى العمري الذي يستحقه. ومن ذلك مثلاً أن اختبار المفردات يطبق ابتداء من المستوى العمري 6 سنوات حتى مستوى راشد المتفوق (الثالث) اعتماداً على عدد الكلمات التي يستطيع المفحوص تحديدها وتعرفها.

والواقع أن الأسئلة التي ينجح فيها المفحوص أو يفشل تظهر مقداراً من اللتشت في المستويات العمرية المتتالية. فنحن لا نجد مفحوصاً يجيب على جميع أسئلة عمره العقلي أو أدنى منه ويفشل في جميع الاختبارات الأعلى من

هذا المستوى؛ وإنما نجد في العادة أن الاختبارات التي ينجح فيها تنتشر في مستويات عمرية عديدة يحددها العمر القاعدي من ناحية والحد الأعلى للعمر من ناحية أخرى. ويحسب العمر العقلي في هذه الحالة بالعمر القاعدي مضافاً إليه أي شهور عقلية لأي اختبار اجتازه من مستويات أعلى من عمره القاعدي.

تفسير المعايير

من أهم مزايَا مقياس ستانفورد بينيه ما يتوافر حوله من بيانات تفسيرية وخبرة إكلينيكية تراكمت طوال نصف قرن. وقد أصبح هذا المقياس عند كثير من علماء النفس الإكلينكيين والتربويين وغيرهم من المهتمين بالتقويم للنسبي مرافقاً للذكاء. وأصبح تحديد مستويات الذكاء يعتمد في جوهره على نسب الذكاء كما يحددها المقياس، رغم ما في هذا الرأي من غلو لا يتسع المقام لمناقشته.

والواقع أن شيوع تصنيف مستويات نسب الذكاء، رغم فائدته في تقنين تفسير الأداء الاختباري، يتضمن بعض المخاطر. فهو كغيره من صور التصنيف للسلوكي يجب أن يستخدم بمرونة كافية، كما يجب ألا يستبعد للبيانات الأخرى التي تتوافر عن المفحوص. وبالطبع لا يوجد حد فاصل تماماً بين فئات التصنيف. ومن ذلك مثلاً لا يوجد خط يميز تماماً بين التخلف العقلي والحالات الهامشية، أو بين التثاق والتثاق الشديد. فقد لوحظ مثلاً أن بعض المفحوصين من ذوي نسب الذكاء 60 يتكيفون تكيفاً عادياً لمطالب الحياة اليومية، بينما نجد بعض الأشخاص من ذوي نسب الذكاء 100 يتطلبون الإيداع في المؤسسات. كما، بعض الأشخاص من ذوي نسب الذكاء 160 قد لا يعيشون حياة متميزة، بينما نجد البعض الآخر من ذوي نسب الذكاء الأقرب إلى 100 يسهمون بعض الإسهامات الممتازة. ومعنى ذلك أن القرارات التي تتحدد حول تشخيص

الضعف العقلي أو التفوق العقلي يجب ألا تعتمد على نسب الذكاء وحدها، وإنما يجب الاهتمام في حالة للضعف العقلي بمتغيرات أخرى مثل النضج الاجتماعي والتوافق الانفعالي والظروف الصحية والجسمية والأحوال الأسرية والظروف المدرسية وغير ذلك من العوامل التي تؤثر في الأداء العقلي. وبالمثل فإن نسب الذكاء للعالية لا ترادف العبقرية، لأن الإيجاز الرفيع قد يتطلب بالإضافة إلى ذلك الابتكار والمواهب الخاصة والمثابرة والاهتمام بالهدف وغيرها من العوامل العقلية والانفعالية والدافعية.

ثبات المقياس

لجريت دراسات للثبات على طبعة 1937 من المقياس باستخدام طريقة الصور المتكافئة على فترات زمنية مقدارها أسبوع أو أقل (للتحقق من مصدر تبين الخطأ: الاستقرار الزمني وتكافؤ المحتوى). وقد لوحظ أن المقياس يميل إلى أن يكون أكثر ثباتاً في الأعمار الكبيرة منه في الأعمار الصغيرة، وفي مستويات نسب الذكاء الدنيا عنه في نسب الذكاء العليا.

والواقع أن زيادة للثبات مع الزيادة في العمر هي من خصائص الاختبارات بوجه عام، وتنتج جزئياً عن التحكم الأفضل في ظروف إعطاء الاختبار في حالة المفحوصين للكبار من (وبخاصة حين نقارنهم بأطفال ما قبل المدرسة). والعوامل الأخرى المساهمة في زيادة معاملات الثبات بطء معدل النمو مع التقدم في العمر.

لما عن ارتفاع معاملات الثبات في حالة نسب الذكاء الدنيا في أي عمر زمني فترتبط فيما يبدو بخصائص مقياس ستانفورد بيليه على وجه الخصوص.

فقد أشرنا إلى أن الأوزان النسبية لأمتلة الأعمار المختلفة تختلف، فهي في الأعمار الدنيا مقدارها شهر، وفي الأعمار المتوسطة شهران، وفي المستويات العليا قد تكون 4 شهور أو 5 شهور أو 6 شهور. ومعنى ذلك أن طريقة وزن الدرجات ميل إلى تضخيم أخطاء القياس في المستويات العليا، لأن النجاح أو الفشل للناجم عن الصدفة في سؤال واحد يؤدي إلى فرق أكبر في الدرجة الكلية في هذه المستويات عنه في المستويات الدنيا. وحيث أن المفحوصين من ذوي نسب الذكاء العالية (من أي عصر زمني) عادة ما يختبرون في المستويات العليا من المقياس، فإن نسب نكائهم نتيجة لذلك تتضمن مقداراً أكبر من أخطاء المقياس وبالتالي تتخفص معاملات الثبات.

وقد حسبت معاملات الثبات لفئات مختلفة من مدى نسبة الذكاء في طبعة 1960، لأنه لوحظ أن جنول انتشار درجات الصورئين (م، ل) من طبعة 1937 يأخذ شكل المروحة، أي أن مدى التشتت يتضاقل كثيراً في المستويات للدنيا، أما المستويات العليا فتظهر تشتتات واختلافات واسعة.

ويوجه عام يمكن القول أن مقياس ستانفرد بينيه على درجة عالية من الثبات، بالرغم من تجانس مجموعات عينة التقين من حيث العمر للزماني، مما يؤثر بالطبع في معاملات الارتباط عموماً ويؤدي إلى نقصانها. وقد تزلوحت هذه المعاملات بالنسبة للأعمار $2 \frac{1}{2}$ - 5 83 (ونسب الذكاء 140-149) 0.91 (النسب الذكاء 60-69). وبالنسبة للأعمار من 6-3 تزلوحت بين 0.91، 97، ولفص المستويات من نسب الذكاء، وبالنسبة للأعمار من 14-18 كانت معاملات الثبات المطابقة لمستويين هي 0.95، 0.98.

صدق المقياس

تتوافر بيانات كافية عن الصدق المرتبط بالمحككات سواء كان صدقاً تلامزياً أو صدقاً تنبئياً في ضوء محك للتحصيل المدرسي. فمنذ ظهر مقياس ستانفورد بينيه عام 1916 لم يتوقف الباحثون عن حساب معاملات الارتباطات بينه وبين المدرجات المدرسية وتقديرات المعلمين ودرجات الاختبارات التحصيلية المقننة ويتراوح معظم هذه المعاملات بين 0.40، 0.75 كما ارتبط بالنظم والتخلف المدرسي بوجه عام.

وقد وجد أن هذا المقياس، كغيره من مقاييس الذكاء، يرتبط ارتباطاً عالياً بالأداء في جميع المواد الدراسية الأكاديمية تقريباً، إلا أن ارتباطه أعلى ما يكون بالمقررات اللغوية (كمواد اللغات أو المواد الاجتماعية). وقد تأكدت هذه النتائج عند اختبار التلاميذ والطلاب في مختلف المراحل التعليمية، وبالطبع فإن معاملات الارتباط تميل إلى التناقص مع ارتفاع السلم التعليمي بسبب عوامل الانتقاء وعدم ملائمة سقف الاختبار مما يؤدي إلى نقصان التباين بين المفحوصين.

وفي تفسير نسبة الذكاء يجب أن نؤكد أن مقياس ستانفورد بينيه، كغيره من اختبارات الذكاء العام، هو في الأغلب مقياس للاستعداد المدرسي أو الأكاديمي العام، ومشبع إلى حد كبير بالمحتوى اللفظي، وبخاصة في اختبارات المستويات العمرية العليا. ولذلك فإن الأفراد من ذوي النقص اللغوي، أو أولئك الذين تتوافر لديهم قدرات غير لغوية يحصلون على درجات منخفضة نسبياً في المقياس. وكذلك فإنه توجد بعض المجالات لا يلعب فيها الاستعداد المدرسي أو الفهم اللغوي دوراً هاماً، وفيها لا تصلح اختبارات الذكاء من نمط ستانفورد بينيه.

ومعنى ذلك أن استخدام الاختبارات - بصفة عامة - يجب أن يكون في المواقف التي نلتمها، وإلا فإن الباحث يتوقع من المقياس ما لا يعطيه.

صدق التكوين الفرضي

من أهم المحركات التي استخدمت في انتقاء اختبارات مقياس ستانفورد بينيه محك تمايز العمر. ومعنى ذلك أن المقياس يقيس قدرات متزايدة مع العمر منذ الطفولة حتى الرشد (في ظل إطار ثقافي وحضاري معين بالطبع).

وقد استخدم أيضاً محك الاتساق الداخلي في انتقاء الأسئلة. وقد أكدت النتائج وجود نوع من التجانس الوظيفي في مقياس ستانفورد بينيه، بالرغم من التتبع الظاهري للمحتوى، والدليل على هذا أن متوسط معاملات الارتباط بين السؤال والمقياس هو 0.66 في طبعة عام 1960 إلا أن من قواضع طلبة الطبع اللغظي على المقياس، ويتضح هذا من ارتفاع معامل الارتباط بين الأسئلة اللغظية والأداء الكلي في المقياس عنه في حالة الأسئلة غير اللغظية.

ومن البراهين الإضافية على صدق التكوين الفرضي نتائج الدراسات العلمية العديدة التي أجريت على اختبارات مقياس ستانفورد بينيه. ويعتمد هذا البرهان على أنه لو كانت نسب الذكاء قابلة للمقارنة في مختلف الأعمار، فإن المقياس يجب أن يكون له نفس التكوين العملي في جميع مستويات الأعمار، بل يجب أن يكون المقياس مشعباً بعامل واحد حتى يمكن تفسير نسبة الذكاء تفسيراً واضحاً. لما إذا كان للمقياس مشعباً بعاملين طائفتين أو أكثر فإن نسبة الذكاء للواحدة التي يحصل عليها الفرد متحدون (نسبة ذكاء 120 مثلاً) قد تدل على قدرة معينة في بعض الحالات، وقدرة أخرى في البعض الآخر.

وقد قام ماكتمر بإجراء عدة تحليلات عاملية منفصلة لاختبارات مقياس ستانفورد بينيه في 14 مستوى عمري (2، 2½، 3، 3½، 4، 4½، 5، 6، 7، 8، 9، 11، 13، 15، 18). وكان عدد الأفراد المستخدم في كل تحليل عاملي يتراوح بين 99، 200، كما تراوح عدد الاختبارات بين 19، 35. وفي كل الأحوال حسبت معاملات الارتباط الرباعية بين المفردات، وخضعت مصفوفات الارتباط للتحليل العاملي. وتؤكد نتائج ماكتمر أن الأداء في مقياس ستانفورد بينيه يمكن تفسيره في ضوء عامل عام واحد، كما وجدت عوامل طائفية إضافية في قليل من المستويات العمرية. إلا أن إسهام هذه العوامل في التباين كان قليلاً. كما ثبت أن العامل العام المشترك الذي وجد في المستويات العمرية للمقاربة كان متشابهاً إلى حد كبير. ومن أهم للنتائج التي توصل إليها ماكتمر أن العامل يصبح أكثر لفظية كلما تقدمت الأعمار. ومن ذلك مثلاً أن تشبع العامل للعام باختبار المفردات ارتفع من 0.59 في سن 6 سنوات على 0.91 في سن 18 سنة.

أجرى جونز دراستين عامليتين منفصلتين لاختبارات مقياس ستانفورد بينيه على 4 مجموعات من الأطفال أعمارهم (7، 9، 11، 13) على التوالي، وكانت كل مجموعة تتكون من 100 ذكر، 100 أنثى. ووجد أنه في كل عمر يوجد عدد من القدرات المتميزة ولكنها مرتبطة، ومنها العوامل اللفظية والتذكرية الاستدلالية والمكانية والإدراكية. وقد اختلفت هذه العوامل أوزانها للنسبة من مستوى عمري لآخر.

والواقع أن الأساليب الإحصائية التي استخدمها ماكتمر وجونز من النوع الذي يؤكد دور العامل العام في الحالة الأولى (دراسات ماكتمر)، ودور العوامل

الطائفية المختلفة في الدراسات التي قام بها جونز. وحين نتأمل النتائج التي توصل إليها الباحثان نجد أنها تؤكد وجود قدر مشترك في القياس ككل، وهي خاصية من خصائص بناء المقياس حسب طريقة انتقاء أسئلته التي كانت تؤكد ارتفاع معاملات الارتباط بين السؤال والدرجة الكلية في المقياس. وبالإضافة إلى ذلك فإن الأداء يتأثر كذلك بعدد من للقرات الخاصة التي يختلف تكوينها تبعاً لمستوى العمر الذي تختبره.

أما عن الدراسات التي أجريت في مصر لتحديد المعالم السيكومترية لمقياس ستانفورد بينيه فإنها قليلة. وتتلخص في دراسة قام بها الدكتور لويس كامل مليكة عام 1960 للتحقيق من صدق الترتيب المبدئي لاختبارات المفردات في المقياس، وذلك بتطبيقها على مجموعات كبيرة من الأفراد ومن مختلف فئات السن (بين 3 سنوات، 30 سنة) بلغ عددها 1642 مفحوصاً وقد تكون الدراسة الوحيدة التي تناولت مقياس ستانفورد بينيه ككل بحث محمود الهولاري عام 1963 عن ثبات المقياس وصدقته. وقد استخدم في حساب الثبات طريقة إعادة الاختبار على 95 تلميذاً بالسنه الثالثة الإعدادية (متوسط أعمارهم 14 عاماً) بفواصل زمني بين شهر وشهرين وبلغ معامل الارتباط 0.83 أما عينة للصدق فتألفت من 160 تلميذاً بالسنه الثالثة الإعدادية أيضاً، وقد حسب لهم معاملات الارتباط بين درجات المقياس (نسبة الذكاء) وحالات الرسوب والنجاح في امتحان الشهادة الإعدادية من ناحية، والمجموع الكلي للدرجة الامتحان من ناحية ثانية، ودرجات مواد هذا الامتحان من ناحية ثالثة. وقد وجد هذا الباحث أن معامل الارتباط اللتائي بين نسب الذكاء وحالات النجاح والرسوب هو 0.54، ومعامل الارتباط بين نسب الذكاء والمجموع الكلي لامتحان الشهادة الإعدادية 0.818 أما معاملات ارتباط المقياس بدرجات المولد الدراسية فكانت على النحو الآتي:

الصحة والعلوم	0.596	الجغرافيا	0.509	اللغة العربية	0.534
التربية الوطنية	0.511	اللغة الإنجليزية	0.485	الفرنسية	0.489
التاريخ	0.487	الجبر	0.511	الحساب	0.557

وأعد الباحث بعد ذلك مصفوفة معاملات ارتباط درجات الذكاء ودرجات المواد الدراسية المختلفة وأخضعها للتحليل العدلي فوجد أن تتبع مقياس ستانفورد بينيه بالعامل العام 0.784 وهو أعلى للتبعيات بهذا العامل وفي هذا برهان مبدئي على صدق التكوين للفرضي للمقياس، وجد أن باقي المتغيرات كانت درجات التحصيل المدرسي، ولم تتضمن المصفوفة أي مقياس آخر للذكاء.

مقاييس وكسلر للذكاء

ظهرت الطبعة الأولى من مقاييس وكسلر والتي عرفت باسم مقياس وكسلر بلغوي للذكاء عام 1929. وكان أحد الأغراض الهامة من إعداد هذه المقاييس توفير اختبار الذكاء يصلح للراشدين. وقد أشار وكسلر في ذلك الحين إلى أن الاختبارات التي كانت متاحة حتى ذلك الوقت، ومنها بالطبع مقياس ستانفورد بينيه، محدودة في جوهرها للأطفال وتلاميذ المدارس، وكانت تستخدم مع المراهقين والراشدين عن طريق إضافة بعض الأسئلة الأكثر صعوبة من نفس النوع. ولذلك فإن من المشكلات الهامة في هذه الاختبارات أن محتواها لا يثير اهتمام الراشدين وإذا لم يتوافر في المقياس حد أدنى من الصدق الظاهري يصعب تكوين علاقة طيبة بين القاحص والمفحوص الراشد.

وبالإضافة إلى ذلك فإن المبالغة في تأكيد عامل السرعة في معظم الاختبارات الفردية يؤدي إلى تعطيل أداء الراشدين. وكذلك فإن الاهتمام بالتناول

الروتيني للنواحي اللفظية تحظى بالاهتمام الأكبر في هذه المقاييس. ثم إنه يستحيل استخدام معيار العمر العقلي مع الراشدين كما لشرنا. وأخيراً فإن عينات تقنين الاختبارات الفردية مثل مقياس ستانفورد لم تتضمن إلا قليلاً من الراشدين.

ولمواجهة هذه المشكلات أعد وكسلر مقياسه للذكاء، ومن أشهرها مقياسه لذكاء الراشدين. والذي يسمى في الوقت الحاضر Wechsler Adult Intelligence Scale WAIS. والذي ظهر في عام 1955 محل مقياس وكسلر بغير التقديم. وبالطبع يتشابه المقياس الجديد في الشكل والمحتوى مع المقياس السابق. إلا أنه تغلب على بعض الصعوبات الفنية الخاصة بمدى تمثيل عينة التقنين. وثبات الاختبارات الفرعية التي يتكون منها المقياس.

وصف مقياس وكسلر لذكاء الراشدين

يتكون مقياس WAIS من 11 اختباراً فرعياً، ستة منها تصنف في مقياس لفظي، 5 منها تصنف في مقياس أدائي أو عملي performance. و فيما يلي وصف موجز لهذه الاختبارات الفرعية مع بيان عدد الأسئلة التي يتكون منها كل اختبار في الطبعتين العربية والأجنبية.

المقياس اللفظي:

يتكون المقياس اللفظي من 6 اختبارات هي:

- 1- المعلومات العامة: information، و يتكون من 29 سؤالاً في الطبعة الأجنبية (26 سؤالاً في الطبعة العربية منها سؤال تمهيدي) تشمل مدى

واسعا من المعلومات العامة التي يفترض شيوعها في ثقافة الراشدين.

وقد روعي تحرر الاختبار من المعارف التخصصية و الأكاديمية.

2- للفهم العام: Comphrnsion ، و يتكون من 14 سؤالا في الطبعة

الأجنبية (10 أسئلة في الطبعة العربية)، و في كل سؤال يشرح

المفحوص ما يجب عمله في ظروف وعينة. ويتطلب هذا الاختبار قدرة

على الحكم العملي و الفهم العام common sense، وهو يشبه في هذا

أسئلة الفهم في مقياس ستانفورد-بينيه، إلا أن محتواه أقرب إلى

اهتمامات الراشدين و نشاطهم.

3- الاستدلال الحسابي: arithmetic، و يتكون من 14 مسألة حسابية في

الطبعة الأجنبية (10 مسائل في الطبعة العربية) تتفق مع مستوى مادة

الحساب في المرحلة الابتدائية. و تعرض كل مسألة شفويا، و يحلها

المفحوص شفويا كذلك دون استخدام الورقة و القلم، و لكل مسألة زمن

محدد.

4- المتشابهات: Similarities، و يتكون من 13 سؤالا في الطبعة

الأجنبية (13 سؤالا في الطبعة العربية) تطلب من المفحوص أن

يحدد كيف يتشابه شيئان (البرقاقة و الموزة مثلا) .

5- مدى الأرقام (أو إعادة الأرقام) : digit span، و فيه تعرض على

المفحوص قوائم تتكون من 3 أرقام إلى 9 أرقام، و يكون عليه إعادتها

شفويا، و في القسم الثاني من الاختبار يكون على المفحوص إعادة

الأرقام بالعكس.

6- للمفردات: vocabulary، و فيه تعرض على المفحوص شفويا و

بصريا في وقت واحد 42 كلمة متزايدة الصعوبة، و يطلب منه تحديد

معنى كل كلمة.

المقياس الأدائي أو العمل:

يتكون المقياس العملي من 5 اختبارات هي:

- 1- رموز الأرقام: digit symbol، و هو عبارة عن صورة من اختبار التعويض الشفري code substitution المألوف في مقاييس الذكاء غير اللغوية. و يتكون مقياس الشفرة من 9 رموز تزأوج مع الأرقام التسعة. و يعطى هذا المفتاح المفحوص و يكون عليه أن يكتب أكبر عدد من الرموز التي تدل على الأعداد الموجودة في ورقة الإجابة في زمن لا يتجاوز 1.5 دقيقة.
- 2- تكميل الصور: Picture completion، و يتكون من 21 بطاقة في الطبعة الأجنبية (15 بطاقة في الطبعة العربية) كل منها يحتوي على صورة ناقصة، و يكون على المفحوص أن يذكر الجزء الناقص من كل صورة.
- 3- رسوم المكعبات: block design، و تشبه مواد هذا الاختبار ما يوجد في اختبار Kohs Block Design. و يتكون من صندوق به 16 مكعباً، و 9 بطاقات على كل منها رسم مختلف (اثنان من هذه للبطاقات للتدريب). وقد لونت ألوجه المكعبات بالألوان الآتية في الأصل الأجنبي: الأحمر فقط و الأبيض فقط و الأبيض_الأحمر معا، أما في الطبعة العربية فهي ملونة بالأزرق فقط و الأبيض فقط و الأصفر فقط و الأبيض_الأحمر معا والأزرق_الأصفر معا. و يطلب من المفحوص إنتاج رسوم على النحو الذي تحده للبطاقات ذات مستويات

متزايدة التعداد تطلب ما بين 4 مكعبات، 9 مكعبات. و لكل بطاقة زمن محدد.

4- ترتيب الصور: Picture arrangement، و يتكون كل سؤال فيه من مجموعة من البطاقات، عليها صور، و يطلب من المفحوص ترتيبها بالتتابع الصحيح بحيث تعطى قصة مفهومة. و لكل سؤال وقت محدد (عدد الأسئلة في الطبعة الأجنبية 8 وفي الطبعة العربية 6).

5- تجميع الأشياء: Object assembly، و يتبع هذا الاختبار نفس النمط المستخدم في اختبار Pinter_Paterson الأدائي. و يتكون في الطبعة العربية من نماذج من الخشب لثلاثة أشياء هي الصبي أو المانيكان، الوجه أو البروفيل، اليد، قطع كل منها إلى قطع مختلفة، و يطلب من المفحوص في كل منها جمع للقطع بحيث تكون الشكل للكمال.

تصحيح المقاييس:

تصحيح اختبارات الاستدلال الحسابي و رسوم للمكعبات و ترتيب الصور و تجميع الأشياء تبعاً لسرعة الأداء و دقته على النحو الآتي:

1- الاستدلال الحسابي: تعطى درجة واحدة لكل مسألة تحل حلاً صحيحاً في الوقت المحدد لها، وتعطى درجة واحدة إضافية للزمن في المسألتين 9، 10 إذا حلت كل منهما في حدود 40 ثانية (الزمن العادي لكل منهما 120 ثانية)، وتعطى درجتان إضافيتان إذا حلت المسألة في حدود 15 ثانية، بحسب الزمن بعد انتهاء المفحوص من قراءة المسألة (النهاية العظمى 14).

2- رسوم المكعبات: تقدر الرسوم حسب الدقة والزمن الأقصى، وتعطى درجات إضافية في حالة إتمام الرسم في زمن أقل. وقد أعد جدول لإعطاء تقديرات حسب زمن الإنجاز (النهاية العظمى 42).

3- ترتيب الصور: تقدر الدرجة في هذا الاختبار على أساس دقة الترتيب والزمن المستغرق فيه، وأعدت جداول لهذا الغرض بحيث تصحح كل من المجموعات الثلاث الأولى على أساس التقدير (صواب أو خطأ)، أما المجموعات الثلاث الأخيرة فتعطى تقديرات جزئية على كل ترتيب لا يطابق الترتيب الصحيح، ولكن يمكن قبوله. وبالنسبة للمجموعتين الأخيرتين تعطى تقديرات إضافية على الترتيب الصحيح إذا تم في زمن معين. وتحسب الدرجة النهائية في هذا الاختبار بجمع التقديرات الجزئية في المجموعات المختلفة (والنهاية العظمى 21).

4- تجميع الأشياء: يصبح الشكل الأول (الصبي) حسب الدقة فقط، أما الشكلان الآخران (الوجه واليد) فيصحح كل منهما حسب الدقة والزمن معاً. وفي الحالة الأخيرة تكون الدرجة في مواد الاختبار هي المجموع الكلي للتقديرات على الدقة والزمن جميعاً. وتوجد بكراسة التعليمات بروتوكولات خاصة لتصحيح كل شكل حسب الدقة، وجدول التصحيح على أساس الزمن (النهاية العظمى 26).

أما لاختبارات المعلومات العامة، ومدى الأرقام (أو إعادة الأرقام)، وتكميل الصور، ورموز الأرقام فإن الدرجة الكلية فيها هي مجموع الإجابات الصحيحة حسب مفتاح التصحيح المبين كراسة التعليمات، والنهايات العظمى لهذه الاختبارات الأربع على التوالي هي: 25، 17، 15، 67.

لما اختبارات الفهم العام، والمتشابهات، والمفردات، فقد قام الدكتور لويس كامل مليكة بإعداد نماذج التصحيح الخاصة بهذه الاختبارات الثلاث ونشرها في كراسة منفصلة عام 1960، وفي هذه النماذج تعطى الأسئلة اختباري الفهم العام والمتشابهات درجتان أو درجة واحدة أو صفر حسب درجة التعميم في الإجابة ونوعه، وهكذا فإن النهاية العظمى لكل من هذين الاختبارين هي: 20، 24 على التوالي. أما اختبار المفردات فإن أسئلته تعطى في نماذج التصحيح درجة واحدة أو نصف درجة أو صفر (النهاية العظمى 42).

المعايير وإجراءات التصحيح

لقد كانت عينة التفتين المستخدمة في مقياس وكسلر للراشدين من العينات للقلائل التي يتوفر فيها شرط التمثيل، فقد كانت تتألف من 1700 شخصاً من الرجال والنساء موزعين في مستويات عمرية سبع بين 16 سنة و 64 سنة وأضيف إلى كل مستوى عمري رجل واحد وامرأة واحدة من المودعين بمؤسسات ضعاف العقول، وكذلك عينة من الممنين عددها 475 شخصاً تزيد أعمارهم على 60 سنة.

وقد حولت الدرجات الخام في الاختبارات الفرعية إلى درجات معيارية متوسطها 10 وانحرافها المعياري 3. وكانت هذه الدرجات الموزونة مشتقة من جماعة مرجعية عددها 500 حالة شملت جميع المفحوصين من سن 20 سنة، 34 سنة في عينة التفتين. وهكذا أصبحت الدرجات الفرعية في صورة وحدات قابلة للمقارنة، ثم حسبت درجات المقياس اللفظي، والمقياس الأدائي، والمقياس الكلي بجمع الدرجات الموزونة في الاختبارات اللفظية الست، والاختبارات الأدائية الخمس، والاختبارات الأحد عشر جميعاً على التوالي. ثم حولت هذه

الدرجات الثلاثة إلى نسبة زكاء الحرافية متوسطها 100 وانحرافها المعياري 15 في كل مجموعة عمرية، وهكذا يتحدد الوضع النسبي للمفحوص بمقارنة بالأشخاص من نفس مستواه العمري.

وقد استخدمت نفس الإجراءات في تقنين الطبعة العربية حيث استخدمت في إعداد الدرجات الموزونة ثلاث فئات عمرية هي:

- (1) من 20 إلى أقل من 25 سنة وعدد أفرادها 114.
- (2) من 25 إلى أقل من 30 سنة وعدد أفرادها 112.
- (3) من 30 إلى أقل من 35 سنة وعدد أفرادها 102.

وقد استخدمت هذه العينة البالغ عددها 328 مفحوصاً في إعداد نماذج ومعايير التصحيح لاختبارات الفهم والمتشابهات والمفردات بالإضافة إلى إعداد جدول الدرجات الموزونة وجدول للنسب للذكاء الانحرافية لكل فئة من فئات العن.

ولتصحيح هذا القياس تعطى المفحوص درجات ثلاث. نسبة للذكاء الكلية ونسبة للذكاء اللفظية ونسبة للذكاء العملية. ويعطى في الطبعة الأجنبية درجة رابعة تدل على ما يسميه وكسلر " معامل الكفاءة " أو (E.Q) ويتحدد بتقدير نسبة زكاء الفرد في الفئة من 20 إلى 24 (في اختبار وكسلر بغير طبعة 1939) وفي الفئة من 29-75، بصرف النظر عن العمر الحقيقي بالمفحوص. ويعتمد هذا المعامل على أساس أن لفئة من 20-24 ولفئة من 25-29 هما اللتان وصلتا إلى أعلى مستويات الأداء في المقياس. ومعنى ذلك أن معامل الكفاءة طريقة للحكم على أداء الفرد في ضوء متوسط أداء الأفراد من المستوى العمري الذي يصل فيه الأداء الاختباري لقضاء.

وينكر الدكتور لويس كامل مليكة أنه يمكن الحصول على " معاماً كفاءة " بصورة مؤقتة عن طريق استخدام جدول نسب الذكاء لفئة من 20، 2 من العينة المصرية آنفة الذكر، لأن هذه الفئة حققت بالفعل أعلى مستوى من الأداء في المقياس.

أسير المعايير

أحد وكسلر عام 1944 فئات الذكاء التي يمكن أن تستخدم في أغراض تصنيف الإكليليكى، وكان أساس التصنيف هو حدود الخطأ المعياري المتوسط الشخص الذي تكون درجته أقل من المتوسط بمقدار يبلغ ثلاثة أمثال الخطأ المعياري أو أكثر بعد من فئة ضعاف العقول، أما الذي يحصل على درجة تقع بين خطين معياريين وثلاثة أخطاء معيارية فيعد من قبيل الحالات الهامشية هكذا. وبين الجدول الآتي أسس تصنيف مستويات الذكاء في مقياس وكسلر.

جدول تصنيف مستويات الذكاء في مقياس وكسلر

التصنيف	جدول الخطأ المعياري	حدود نسبة الذكاء الاعتراف	النسبة المئوية في السكان
ضعاف العقول	3- فأقل	65 فأقل	2.2
الحالات الهامشية	من 2- حتى 3-	من 66 إلى 79	6.7
الأغباء	من 1- حتى 2-	من 80 حتى 90	16.1
المتوسطون	من 1- حتى 1+	من 91 حتى 110	50.0
أعلى من المتوسط	من 1+ حتى 2+	من 111 حتى 119	16.1
الممتازون	من 2+ حتى 3+	من 120 حتى 127	6.7
المتأزرون جداً	من 3+ فأكثر	من 128 فأكثر	2.2

ثبات المقياس

حسبت معاملات الثبات للطبعة الأجنبية من مقياس وكسلر لذكاء الراشدين لفئات الأعمار 18-19، 24-34، 45-54 لختبرت من عينة التقنين على أسس تمثيلها لمدى العمر الذي تشملته هذه العينة. واستخدمت طريقة التجزئة النصفية (بعد تصحيح سبيرمان- براون) في حساب ثبات كل اختبار فرعي فيما عدا اختبائي رموز الأرقام ومدى (إعادة) الأرقام حيث استخدمت معهما طريقة الصور المتكافئة لعدم صلاحية طريقة التجزئة النصفية لهما بسبب غلبة عامل السرعة الشديدة عليهما.

وبلاحظ أن معاملات الثبات للمقياس الكلي قد بلغت في المجموعات الثلاثة 0.97، وبلغت في المقياس اللفظي 0.96، وفي المقياس الأدائي 0.94 ومعنى هذا أن المقاييس الثلاثة على درجة عالية من الثبات في ضوء الاتساق الداخلي.

أما في الطبعة العربية فقد حسب الدكتور فرج عبد القادر طه ثبات المقياس بطريقة إعادة الاختبار وطريقة التجزئة النصفية، وفي الطريقة الأولى أيد تطبيق المقياس على 40 حالة من العمال فتراوحت معاملات الثبات بين 0.6، 0.9 بالنسبة لكل من نسب الذكاء والاختبارات الفرعية. أما طريقة التجزئة النصفية فقد حسبت لدرجات 70 عاملاً وتراوحت معاملات الثبات للنتائج عنها بين 0.5، 0.9 لنسب الذكاء والاختبارات الفرعية. ولم يوضح الباحث كيف استخدم طريقة التجزئة النصفية لاختبائي رموز الأرقام وإعادة الأرقام بالرغم من عدم صلاحية هذه الطريقة لهما كما أشرنا.

صدق المقياس

يفضل وكسلر في تناوله لصدق مقياسه أن يؤكد " صدق المحتوى ". ومن رأيه أن الاختبارات الأحد عشر المتضمنة فيه تتفق مع تعريف الذكاء، وأن لاختبارات مماثلة استخدمت في مقاييس الذكاء السابقة، وثبتت جدواها في الأغراض الإكلينيكية. ومع ذلك فإنه ينكر بعض البيانات المحدودة عن الصدق التلازمي للمقياس. ومن ذلك المقارنة بين المجموعات التطيمية والمهنية المختلفة، بالإضافة إلى حساب معاملات الارتباط مع تقديرات الأداء على العمل والتقديرات المدرسية، وكانت أغلب الفروق بين المجموعات في الاتجاه المتوقع، ومن ذلك أن درجات العمال أعلى في الاختبارات العملية، وأن درجات الأفراد في المهن غير العملية أعلى في الاختبارات اللفظية. أما عن معاملات الارتباط بالمحكات المهنية والمدرسية فقد أعطى المقياس اللفظي رقماً أعلى من المقياس الكلي. أما المقياس الأدائي فقد أعطى أدنى القيم، وفي كل الأحوال كانت معاملات ارتباط المقياس اللفظي أدنى من معاملات مقياس ستانفورد بينيه. أما عن الصدق للتنبؤي للمقياس فقد ثبت أن للمقياس يتنبأ بدرجة مقبولة بالضعف العقلي كما تؤكد بحوث جيرتن وزملائه.

وقد حسبت معاملات الارتباط بين مقياس وكسلر وغيره من الاختبارات والمقاييس المعدة لقياس الذكاء، وكانت النتائج كما يلي:

- 1- معامل الارتباط بين المقياس ومقياس ستانفورد بينيه 0.80 وهو أعلى في المقياس اللفظي منه في المقياس الكلي والمقياس الأدائي.
- 2- معاملات الارتباط بين المقياس واختبارات الذكاء الجماعية تتراوح بين 0.40، 0.80.

3- معامل الارتباط بين المقياس الأدائي واختبار منيسوتا للوحات الأشكال
.72

4- معامل الارتباط بين المقياس الأدائي واختبار رلفن للمصفوفات المتتابة
.0.70

5- معامل الارتباط بين المقياس الأدائي واختبار بيوت للفهم الميكانيكي
.0.35

صدق التكوين الفرضي

أجريت عدة تحليلات علمية للاختبارات الفرعية والتي يتألف منها مقياس وكسلر. ومن أشهر هذه الدراسات بحوث كوهن على مصفوفات ارتباطية لمجموعات عمرية أربعة من عينة للتقنين هي: (من 18-19، ومن 25-35، ومن 45-54، ومن 60-75+). وقد وجد الباحث أن الاختبارات الفرعية جميعاً تشترك في عامل عام واحد يسهم بحوالي 50% من التباين الكلي للبطارية. وبالإضافة إلى ذلك استخرج ثلاثة عوامل طاقفية هي:

1- عامل للفهم اللفظي وتشيع في اختبارات المفردات والمعلومات العامة والفهم العام والمتشابهات.

2- عامل التنظيم الإدراكي وتشيع في رسوم المكعبات وتجميع الأشياء، وهو عبارة عن مركب من عاملي السرعة الإدراكية وللتصور البصري المكاني اللذين أكتنهما البحوث للعملية المرجحة وبخاصة بحوث ثرستون.

3- عامل الذاكرة وتشيع أساساً في اختبارات الاستدلال الحسابي ومدى (إعادة) الأرقام، ويتضمن كلاً من الذاكرة الاستظهارية المباشرة للمواد الجديدة واستدعاء المواد التي سبق تعلمها.

ومن للطريف أن نذكر أن نتائج هذا البحث لم تدعم الإجراء الذي يستخدم في تصحيح هذا المقياس أي تقسمه إلى مقياس لفظي ومقياس عملي أدائي لكل منهما نسبة ذكاء منفصلة. ورغم أن استخدام نسب الذكاء الكلية يدعمه وجود عامل عام مشترك في جميع الاختبارات، فإن عامل الفهم اللفظي لم يظهر إلا في 4 اختبارات لفظية فقط من 6 اختبارات. أما عامل الذاكرة فقد وجد في الاختبارين اللفظيين الباقيين بالإضافة إلى اختبارات فرعية أخرى من المقياسين عند الراشدين الأكبر سناً. أما عامل لتنظيم الإدراكي فكانت تشعباته الدالة في اختبارين فقط من الاختبارات الأدائية الخمس، ولما الاختبارات الأدائية للثلاث الباقية فهي فيما يبدو لها عواملها الخاصة التي تشترك مع الاختبارات الأخرى المتضمنة في بطارية المقياس.

وصف مقياس وكسار لذكاء الأطفال

يتكون مقياس وكسار لذكاء الأطفال من الأسئلة السهلة والتي كان يتألف منها مقياس وكسار بقلو الأصلي. ويقسم المقياس، كما هو الحال في مقياس الراشدين، إلى قسمين: أحدهما لفظي والآخر عملي، والمجموع الكلي للاختبارات الفرعية 12 اختباراً منها اختباران احتياطيان أو إضافيان عدد للضرورة، على النحو الآتي:

القسم اللفظي	القسم العملي (الأدائي)
1- للمعلومات العامة	6- تكلمة الصور
2- الفهم العام	7- ترتيب الصور
3- الاستدلال الحسابي	8- تصميم (رسم) المكعبات
4- المتشابهات	9- تجميع الأشياء
5- المفردات	10- للترميز
(إعادة الأرقام)	(المتاهات)

وقد اعتبر وكسلر الاختبارات الاحتياطية هي تلك التي ارتبطت لدى الارتباطات ببقية المقياس، وعلى ذلك فإن اختبار مدى (إعادة) الأرقام في المقياس اللفظي واختبار المتاهات في المقياس العملي من هذا النوع. ويتطابق اختبار الترميز مع اختبار رموز الأرقام في مقياس الراشدين فيما عدا غلبة الأسئلة السهلة عليه. أما الاختبار الوحيد الذي ليس له نظير مقياس في الراشدين فهو اختبار المتاهات، ويتكون 8 متاهات ورقية متزايدة الصعوبة، وتصحح في ضوء الزمن وعدد الأخطاء وفي رأي وكسلر أن اختباري الترميز والمتاهات يمكن أن يحل أحدهما محل الآخر، ويعتمد ذلك على القرار الذي يتخذه الفاحص. ويتطلب اختبار الترميز زمناً أقل من اختبار المتاهات، وقد يفضل بعض الفاحصين لذلك.

ولا يختلف مقياس الأطفال في إجراءات تصحيحه من مقياس الراشدين إلا اختلافات طفيفة. وتتلخص في تحويل الدرجات الخام في كل اختبار فرعي إلى درجات معيارية اعتدالية في إطار المجموعات العمرية التي ينتمي إليها المفحوص. وقد أعد وكسلر جداول لهذه الدرجات الموزونة لفئات عمرية

مقدارها أربعة شهور ما بين سن 5 سنوات حتى سن 15 سنة. وتشبه للموزونة هذا درجات مقياس الراشدين في أن متوسطها 10 وانحرافها المعياري 3، ثم تجمع الدرجات الموزونة وتحول إلى نسب ذكاء لحرافية متوسطها 100 وانحرافها المعياري 15 سواء للمقياس اللفظي أو للمقياس الأدائي أو للمقياس الكلي. وكذلك اهتم وكسلر باقتراح بعض الطرق الإحصائية لحساب الأعمار العقلية المقابلة لدرجات مقياس الأطفال. ورغم أن هذه الأعمار العقلية ليست من خصائص نسب الذكاء الانحرافية التي تتكلف منها معايير المقياس إلا أنه إذا كان في ذلك يواجه مطلباً عملياً ملحاً نتيجة لمشروع مفهوم العمر العقلي بين خاصية اللام وعلمتهم بالنسبة لميدان قياس للذكاء. وقد أعدت بالفعل جداول لتحويل الدرجات الخام في المقياس إلى أعمار عقلية.

وتتألف عينة ثنتين مقياس وكسلر للأطفال من 100 طفل و 100 طفلة في الأعمار من 5 سنوات إلى 15 سنة، وكان المجموع الكلي للعينة 2200 مفحوصاً كلهم من تلاميذ للمدارس باستثناء 55 حالة من حالة الضعف العقلي اختيرت من المؤسسات.

ثبات المقياس

حصلت معاملات الثبات بطريقة التجزئة للنصفية لكل اختبار فرعي في مقياس وكسلر للأطفال ولكل من المقياس اللفظي أو المقياس العملي، والمقياس الكلي، لعينات $7\frac{1}{2}$ سنة، $10\frac{1}{2}$ سنة، $13\frac{1}{2}$ سنة، واحتوت كل عينة عصرية على 200 مفحوص، ولم تستخدم طريقة التجزئة النصفية مع اختباري الترميز ومدى الأرقام لحدم ملاحظتها (لأسباب وضحتها آنفاً)، ولذلك حسب الثبات لهما بطريقة الصور المتكافئة، وكانت معاملات الثبات للمقياس الكلي للمجموعات

المعربة الثلاثة 0.92، 0.95، 0.94، وعلى التوالي، والمقياس اللفظي: 0.88، 0.99، 0.96 على التوالي، والمقياس الأدائي: 0.85، 0.89، 0.90 على التوالي. وهي معاملات مرتفعة وتدل على ثبات المقياس.

أما معاملات ثبات الاختبارات الفرعية فكانت منخفضة بصفة عامة فيما عدا اختبار المفردات كما أن هذه المعاملات كانت أكثر انخفاضاً في حالة الأطفال الصغار منها في حالة الأطفال الكبار.

صدق المقياس

لا تتضمن كراسات تعليمات مقياس وكسلر للأطفال بيانات عن صدقه، ولو أن بعض الشواهد تدل على ذلك، وخاصة ما لوحظ من أنه من حالة واحدة من بين 55 حالة ضيف عتلي تضمنها عينة للتقنين وصل مستوى نكاتها في المقياس العتلي إلى نسبة ذكاء انحرافية 70، وكان متوسط نسب ذكاء هذه المجموعة 57. كما أثبت عدد من الباحثين الصدق للتلاميذ المقياس باستخدام محكات اختبارات التحصيل الأكاديمي، وبلغت معاملات الارتباط حوالي 0.60، وكانت معاملات ارتباط المقياس اللفظي بهذه المحكات بالطبع أعلى من المقياس العملي.

أما بحوث التحليل العامل التي أجريت على المقياس بهدف تحقيق صدق للتكوين الفرضي والتي يلخصها ليتل أما عام 1960 فقد استخرجت عوامل تشبه ما وجد في تحليل مقياس لاراشدين وهي: العامل العام، والفهم اللفظي، والعامل المكاني الإدراكي، وعامل الذاكرة. ولا يوجد إلا قدر قليل من

التطابق بين عامل الفهم اللفظي والاختبارات الفرعية. المقياس اللفظي، والعامل المكاني الإدراكي والاختبارات الفرعية في المقياس الأدائي العملي.

تكوين المقياس

يمكن القول أن مقياس وكسلر للأطفال مقياس جيد لنكاء الأطفال ويتفق في الخصائص مع كثير من الاختبارات الفرعية. ومع ذلك فمن رأى استلزي أن الأساس المنطقي لهذا المقياس ظاهر التناقض. فقد أشرنا إلى أن مقياس وكسلر بلفيو الأصلي لمقابلة الحاجة إلى اختبار الراشدين التي لا يوفرها مقياس ستانفورد بينيه أو أي مقياس آخر يعتمد على مجرد (رفع مستوى) مقاييس الأطفال. وبهذا حقق وكسلر هذا الهدف سعي إلى إعداد مقاييس للأطفال تعتمد في جوهرها على (خفض مستوى) مقاييس للراشدين، بل أنه صار بالتناقض إلى نهايته حينما أعد مقياساً للأطفال ما قبل المدرسة.

مقياس وكسلر لنكاء أطفال ما قبل المدرسة

ظهر مقياس وكسلر لنكاء أطفال ما قبل المدرسة عام 1967، ويهدف إلى قياس ذكاء الأطفال من سن 4 سنوات إلى $6\frac{1}{2}$ سنة، ويتألف من 11 اختباراً فرعياً يستخدم منها 10 اختبارات فقط في تحديد نسبة الذكاء، ومنها نجد 8 اختبارات عبارة عن " خفض لمستوى " الاختبارات للمناظرة لها في مقياس الأطفال آنف الذكر، أما الاختبارات الثلاثة الباقية فقد أعدت خصيصاً لهذا المقياس. وتصنف الاختبارات الفرعية في هذا المقياس، كما هو الحال في مقياسي Wisc, Wais إلى قسمين: لفظي وعملي، على النحو الآتي (وقد وضعنا علامة * أمام الاختبارات الجديدة).

المقياس الأدائي (العملي)

- * بيت الحيوان
- * تكلمة الصور
- * المناهات
- * الرسوم الهندسية
- رسم المكعبات

المقياس اللفظي

- المعلومات
- المفردات
- الحساب
- * المشابهات
- الفهم
- * الجمل (اختبار إضافي)

ويوصى وكسلر باستخدام الاختبارات اللفظية والعملية بالتناوب لتوفير التنوع والاحتفاظ بميل الطفل وتعاونه. ويتطلب المقياس في تطبيقه ما بين 50 دقيقة و 75 دقيقة في جلسة واحدة أو جلستين اختباريتين.

وقد حل اختبار الجمل في هذا المقياس محل اختبار مدى (إعادة) الأرقام في مقياس Wisc، وفيه يكرر للطفل كل جملة مباشرة بعد عرضها عليه شفويًا. وهذا الاختبار يمكن أن يستخدم كبديل لأي اختبار لفظي آخر، أو يطبق كاختبار إضافي لتزويد الفحص بمعلومات أكثر عن الطفل، وفي كلتا الحالتين لا يستخدم في حساب نسبة الذكاء. أما اختبار بيت الحيوان فيشبه اختبار رموز الأرقام في مقياس Wais واختبار الترميز في مقياس Wisc. ويتضمن اختبار مفتاحاً يعطي للطفل على لوحة وتلّف من صور لكب ودجاجة وسمكة وقطة، لكل منها بيت (أسطواني) ملون بلون مختلف موضوع تحت الصورة، وعلى الطفل أن يضع الأسطوانة (البيت) ذات اللون المناسب في فتحة تحت صورة الحيوان على اللوحة. وتتحدد درجة الطفل في ضوء الزمن وعدد الأخطاء وعدد الاستجابات المحذوفة. أما اختبار "الرسوم الهندسية" فيطلب من الطفل محاكاة 10 رسوم بسيطة بقلم ملون.

اختبارات القدرة المتخصصة

اختبارات القدرة للتخصصة

أشرنا في الفصلين السابقين إلى أن اختبارات الذكاء العام لا تشمل فيما تقيسه ميدان " القدرات العقلية "، وقد ظهر هذا النقص حتى قبل الاهتمام ببطاريات الاستعدادات المتعددة، وتمثل ذلك في ظهور الاختبارات التي تقيس ما كان يسمى " الاستعدادات الخاصة " وأشهرها وأكثرها تذكيراً اختبارات " الاستعداد الميكانيكي "، ثم تزايدت الحاجة إلى هذه الاختبارات التي تقيس ما يمكن أن نسميه " القدرات المتخصصة " بزيادة الاهتمام في علم النفس التطبيقي بمشكلات التوجيه والانتقاء والتوزيع والإرشاد النفسي، فكتائر عددها وشملت مجالات أكثر تنوعاً ومن ذلك ميدان النشاط الموسيقي واللغوي والمهني والكتابي، بالإضافة إلى النشاط الميكانيكي الذي أشرنا إليه.

وقبل أن نتناول نماذج من اختبارات القدرات المتخصصة بشيء من التفصيل لا بد أن نوضح مكانة هذه الفئة من الاختبارات في النظرية الحديثة للقدرات العقلية وبإيجاز نقول أن هذه الاختبارات ظهرت في وقت كان الاهتمام كله مركزاً على ما يسمى " اختبارات الذكاء العام " والتي أشرنا إليها في الفصل الرابع، بهدف أن تكون مكملة لمفهوم " نسبة الذكاء " في وصف النشاط العقلي للمفحوص، إلا أنه مع التقدم في مناهج البحث وفي نظرية القدرات العقلية، وخاصة بعد شيوع استخدام منهج التطويل العامل والنماذج المؤسسة عليه، أصبح من الشائع اعتبار مفهوم " الذكاء " ذاته يتألف من عدد كبير من " القدرات العقلية " يصل في أحد النماذج المعاصرة ونقصد به نموذج جيلفورد إلى أكثر من 120 قدرة وزاد نتيجة لذلك الاهتمام ببطاريات الاستعدادات المتعددة. بل إن

بعض الاختبارات التقليدية " للاستعدادات الخاصة " أصبحت من مكونات بعض بطاريات الاستعدادات كما أشرنا في الفصل الخامس.

والسؤال الذي تطرحه أستاذي هو: ما هو موضع هذه الاختبارات في الوقت الحاضر ؟

الواقع أن لهذه الاختبارات وظيفتين همتين في التقويم النفسي المعاصر هما:

1- توجد بعض مجالات الملوك الإنساني لا تتوفر لها مقاييس جيدة في بطاريات الاستعدادات المتعددة ومن ذلك الإدراك البصري والسمعي والمهارة الحركية والمواهب للموسيقية والفنية والملوك الابتكاري. وقد يكون السبب في ذلك أن المواقف التي تتطلب استخدام هذه الاختبارات على درجة كبيرة من النوعية بحيث لا تكرر إجمالها في بطاريات الاستعدادات المتعددة. ومعنى ذلك أن هذه الاختبارات ستظل فئة متميزة من الاختبارات العقلية حتى تتوفر بطاريات " شاملة " لقدرات الإنسان. وهنا ما نقتبأ حدوثه في المستقبل القريب لعلم القياس النفسي.

2- نستخدم اختبارات القدرات المتخصصة حتى في الأحوال التي توجد نماذج لها في بطاريات الاستعدادات المتعددة، ومن ذلك اختبارات القدرات الميكانيكية أو الكتابية. والسبب في ذلك أن بعض اختبارات القدرات المتخصصة تتوفر لها بيانات ثقلية واقعية، بالإضافة إلى المرونة في استخدام الاختبارات، وشمول الميدان الذي تقيسه. فبدلاً من استخدام اختبار واحد للقدرات الميكانيكية في إحدى البطاريات المتعددة الاختبارات، يمكن للباحث أن يطبق بطارية لاختبارات القدرة.

ويتوالى في الوقت الحاضر عدد هائل من القدرات التي تقيس " القدرات المتخصصة " ابتداء من المستوى الحاسي (السمع والبصر) حتى مستوى التفكير الابتكاري والنشاط المهني. ولا يتسع المقام لتناول هذه الاختبارات بأي نوع من العرض الشامل ولذلك سوف تقتصر على طائفة من هذه القدرات وما يقيسها من اختبارات.

الختبار للقدرات الحركية

ظهرت الاختبارات التي تقيس خصائص الاستجابات الحركية كالسرعة والتأخر منذ فترة طويلة، ويهتم معظمها بالمهارة اليدوية وقليل منها يهتم بحركات الساق وللقدم والتي تتطلبها بعض المهن والأعمال. وقيس بعضها مزيجاً من الاستعدادات الحركية والإدراكية والمكانية والميكانيكية. وفي أغلب الأحوال فإن اختبارات القدرات الحركية تتطلب أجهزة في قياسها رغم ظهور بعض اختبارات الورقة والقلم في أغراض التقويم النفسي الجماعي. وقد أشرنا في الفصل الخامس إلى بطاريتي فلانجان والاستعدادات العامة اللتين تضمنتا بعض الاختبارات الحركية بالجماعية إلا أن الدراسات التي قام بها فلشمان وملتون تؤكد عدم وجود علاقة بين اختبارات الورقة والقلم واختبارات الأجهزة التي تقيس نفس القدرات الحركية، وفيما يلي أمثلة للأجهزة التي تستخدم في قياس للنشاط الحركي:

- 1- مقياس القوة العضلية (الدينامومتر) Dynamometer وتستخدم في قياس قوة قبضة اليد.
- 2- جهاز قياس زمن الرجوع.

3- اختبار Grawford Small Parts Dexterity وقياس المهارة اليدوية البسيطة.

4- اختبار Purdue Pegboard وقياس النشاط الحركي الدقيق والنشاط الحركي الغليظ.

5- اختبار Stomberg Dexterity وقياس حركات اليد والذراع.

الختبارات القدرات الميكانيكية

إن ما يسمى القدرة الميكانيكية ليس قدرة موحدة لا تقبل التحليل، وإنما هي في الواقع مركب من القدرات الحسية والإدراكية والحركية، بالإضافة إلى قدرات إدراك العلاقات المكانية، واكتساب المعلومات الميكانيكية، وفهم العلاقات الميكانيكية.

ولذلك فإن الاختبارات التي تقيس القدرة الميكانيكية إنما تهتم بمستوى من الأداء أعلى مما تتضمنه الاختبارات الحسية والإدراكية (التمييز السمعي والبصري والحركي والتآزر العضلي والمهارة اليدوية).

ويعد اختبارا التجميع Assembly الذي أعده ستوكويست عام 1923 أول اختبار ظهر في هذا الميدان يقيس بقدرة المفحوص على تركيب أجزاء الأجهزة الميكانيكية مثل جرس الدراجة، أو قفل، أو مصيدة غرآن. ويتألف من 3 سلاسل تستخدم مع مستوى عمري يمتد بين الطفولة والرشد.

وقد عدل هذا الاختبار في جامعة مينيسوتا عام 1930 وأصبح يطلق عليه اختبار مينيسوتا للتجميع الميكانيكي وأضيفت إليه أجهزة جديدة. ويصحح هذا الاختبار كسابقه في ضوء معدل الاستجابة ودقتها. وقد أكنت الدراسات التي

أجريت عليه أنه على درجة كافية من التنبؤ بنجاح طلاب للتعليم الفني المبتدئين في أعمال الورشة، كما ارتباط بالنجاح في بعض المهن الميكانيكية.

وفي عام 1930 ظهر اختبار آخر في جامعة مينيسوتا هو اختبار مينيسوتا للعلاقات المكانية ويتألف من سلسلة من 4 لوحات خشبية يتكون كل منها من 57 قطعة خشبية مختلفة الأشكال، وبعضها غير مألوف ويطلب من المفحوص أن يضع هذه القطع أماكنها في فتحات اللوحة الخشبية، وفي رأي كثير من النقاد أن هذا الاختبار يقيس السرعة والدقة في الاستجابة لتفاصيل العلاقات المكانية، وقدرة المفحوص على التعامل مع الأشياء والمواد المحسوسة. وبعبارة أخرى فإنه لا يقيس القدرة على حل المشكلات ذات الطبيعة الميكانيكية، كما لا يصلح لقياس القدرة على معالجة الأشياء الصغيرة بدقة.

وفي عام 1948 ظهر اختبار مينيسوتا للوحات الأشكال الورقية يعرض على المفحوص على ورقة مطبوعة نفس المشكلات التي كانت تعرض عليه في لوحات الأشكال الحقيقية. ففي كل سؤال من أسئلة هذا الاختبار يعرض على المفحوص جزءان أو أكثر من شكل هندسي بحيث لو جمعت معاً بطريقة صحيحة فإنها تكون الشكل الكلي للكمال. وعلى المفحوص تحديد الشكل الذي يمكن أن يجمع مع الأجزاء الأخرى من بين 5 بدائل لاختيارية. وهكذا يقيس هذا الاختبار القدرة المكانية كما تحدثت في البحوث العلمية للقدرة العقلية. وتؤكد نتائج البحوث التي أجريت على هذا الاختبار أنه يرتبط ارتباطاً دالاً بلوغ الأداء الميكانيكي، والنجاح في الرسم الهندسي والهندسة الوصفية. واستخدمت في حساب صدقه محكات مختلفة مثل درجات مقررات الهندسة والمواد الفنية والورش الميكانيكية وتقديرات المشرفين ومقياس الكفاءة الإنتاجية.

وتوجد فئة أخرى من اختبارات القدرة الميكانيكية تقيس المعلومات الميكانيكية، والاستدلال الميكانيكي والفهم الميكانيكي، وهذه جميعاً تتطلب من المفحوص لفه بالآلات والشائعة وبالعلاقات الميكانيكية وبالعلاقات الميكانيكية، إلا أنها لا تقتصر في المفحوص توفر معلومات فنية أكثر مما يمكن أن توفره الحياة اليومية في المجتمعات الصناعية. ومن أشهر هذه الاختبارات تلك التي أدها بينت وزملاؤه والتي تسمى اختبارات الفهم الميكانيكي وتعرض على المفحوص مشكلات ميكانيكية مصورة وعدة حلول لكل منها، وعليه أن يختار الحل الصحيح. ولهذه الاختبارات ثلاثة مستويات من الصعوبة، وتهدف إلى قياس القدرة على فهم المبادئ الفيزيائية والميكانيكية في مواقف بسيطة نسبياً.

ومن الاختبارات الأحدث نسبياً في قياس القدرة الميكانيكية اختبار الدافعية الميكانيكية المسمى Mellanbruch Mechanical Motivation وهو اختبار مصور يتطلب من المفحوص التعرف على أشياء معينة شائعة في المجتمعات الصناعية المتقدمة. والاختبار الثاني هو SRA Meckalnical Aptitdes Test ويتألف من 3 اختبارات فرعية تقيس:

أ- المعلومات الميكانيكية.

ب- إدراك الأشكال والتصور البصري المكاني.

ج- حل المشكلات الحسابية الخاصة بالأعمال الميكانيكية (في صورة جداول ورسوم بيانية).

ويجب أن نؤكد أن هذه الاختبارات التي تقيس الفهم الميكانيكي تتطلب تعديلاً شاملاً عند محاولة استخدامها في البيئة المصرية، وذلك لأنها مشبعة تشبعاً كبيراً بالعناصر الثقافية السائدة في المجتمعات التي صنعت فيها.

اختبار القدرة الكتابية

القدرة الكتابية - كالقدرة الميكانيكية - ليست قدرة بسيطة وإنما هي على درجة من التعقد والتركيب. ورغم أن بعض بطاريات الاستعدادات المتعددة التي تناولناها في الفصل الخامس - وخاصة بطارية الاستعدادات الفارقة وبطارية الاستعدادات العامة - تضمنت اختبارات فرعية لقياس القدرة الكتابية، فإن تراث التقويم النفسي يتضمن طائفة من الاختبارات التي تقيس هذه القدرة على وجه الخصوص، بل توجد بطاريات لقياس القدرة الكتابية، وعموماً نستطيع القول أن اختبارات هذه القدرة تجمعها خصائص مشتركة يوضحها الجدول الآتي الذي يلخص محتوى 6 بطاريات للقدرة الكتابية واختباراتها الفرعية. ومنه يتضح أن البطاريات الشاملة تسعى إلى قياس عدد من العمليات للنوعية، بينما تسعى البطاريات الأقل شمولاً إلى قياس السرعة والدقة الإدراكيين والتي تعد أحد جوانب القدرة الكتابية وليست كلها. وقد يكون من العوامل الكامنة وراء قصر القدرة الكتابية على السرعة الإدراكية ما يؤكدته نتائج البحوث من أهمية القدرات الإدارية بوجه عام فيها. إلا أن " للقدرة الكتابية " تشمل أنواعاً متنوعة من الأعمال تختلف في مدى القدرة ومستواها، ولذلك فإن مقاييسها تتضمن عمليات عقلية أخرى أساسية، بالإضافة إلى بعض القدرات الحسية والمهارات اليدوية عند الضرورة.

البطارية	الاختبارات الفرعية
بطارية ديترويت Detroit	الخط: السرعة والدقة المراجعة: السرعة والدقة الحساب البسيط السرعة والدقة في الأداء الحركي معرفة المصطلحات للتجارية البسيطة ترتيب الصور التصنيف: السرعة والدقة الترتيب الأبجدي
بطارية الاستعدادات الكتابية العامة General Clerical	المزوجة: اكتشاف الأخطاء في الأسماء والأعداد ، الترتيب الأبجدي الحساب: الجمع البسيط الحساب: تحديد أخطاء الجمع المسائل الحسابية البسيطة التهجي فهم القراءة معرفة معاني الكلمات الاستخدام للغوي: القواعد النحوية

الاختبارات الفرعية	البطارية
مقارنة الأعداد مقارنة الأسماء	بطارية مينيسوتا
التهجى الحساب المراجعة: السرعة معرفة معاني الكلمات لنقل وللمسخ: الدقة الاستدلال	بطارية برندو Purdoue
العمليات العددية معاني الكلمات التصنيف	بطارية التوظيف Short Bmployment
المهارات اللفظية المهارات العددية للتعليمات المكتوبة المراجعة: السرعة التصنيف للترتيب الأبجدي	بطارية تورس Turse

بد في المكتبة العربية بطاريتان لقياس القدرة الكتابية هما:

(1) لاختبارات القدرة الكتابية للدكتور محمد عبد السلام أحمد

وتتألف من اختبارين هما:

- اختبار تصنيف الأعداد.

- اختبار تصنيف الأسماء.

اعتمدا على تحليل لأعمال الكتبية قام به الدكتور محمد عبد السلام أحمد بمعاونة ديوان الموظفين عام 1953. ولقن الاختبار الأول على عينة حجمها 822 مفحوصاً من الجنسين من الحاصلين على شهادة الثانوية للعالمية أو دبلوم للتجارة الثانوية تتراوح أعمارهم بين 18، 28 سنة. ولقن الاختبار الثاني على عينة عددها 880 فرداً من الجنسين تتراوح أعمارهم أيضاً بين 18، 28 سنة من حملة الثانوية العامة المتقدمين لشغل الوظائف الكتبية عن طريق إعلانات ديوان الموظفين في ذلك الوقت. وأعدت للاختبارين معايير مئينية وثانية وحسب لهما معامل لثبات بطريقة للتجزئة النصفية فبلغ 0.97 والواقع أن استخدام طريقة للتجزئة النصفية في حساب لختبارات القدرة للكتبية - التي تعتمد اعتماداً كلياً على السرعة - يثير كثيراً من المشكلات أشرنا إليها في الفصول السابقة، إلا إذا كان المؤلف لجأ إلى تجزئة زمن الاختبار، وهذا ما لم توضحه كراسة تعليمات هذه الاختبارات. أما عن صدق الاختبارين فقد حسب مؤلفهما معاملات الارتباط بينهما وبين لختبارات المهن الكتبية والذكاء الاجتماعي فلو حظ أن الاختبارين يشتركان مع الاختبارات التي تقيس الجوانب الأكبية: السرعة، الدقة، سهولة استخدام اللغة، الألفة بالأعداد، الذاكرة المباشرة، المهارة اليدوية.

(2) لختبارات المهن الكتبية من إعداد الدكتور محمد عماد الدين إسماعيل

والدكتور سيد عبد الحميد مرسى:

وتقيس مكونات ثلاث هي:

- القدرة الحدية.

- السرعة والدقة،
- الاستدلال اللغوي.

وقد حسب ثبات الاختبارات الثلاث بطريقة التجزئة للصفحة فبلغت 0.86، 0.84، 0.85 على التوالي. ومرة أخرى نشير مسألة استخدام الطريقة الملائمة لحساب ثبات الاختبارات التي تعتمد على السرعة. واستخدم في حساب صدق الاختبارات تقديرات مادة المحاسبة والترجمة وتقديرات المدرسين والمشرفين وتراوحت معاملات الارتباط بين 0.42، 0.54 وحسبت المعايير المئينية للاختبارات.

اختبارات القدرة الموسيقية

وجد في الوقت الحاضر عدد قليل من الاختبارات التي تصلح لتقاييس القدرة الموسيقية؛ قد يكون أهمها ظهور أو أكثرها أهمية لاختبارات سي شور التي تسمى: Seashore Measures of Musical Talent.

وقد ظهرت في صورتها الأصلية عام 1919، واستمرت شائعة الاستخدام في الولايات المتحدة دون تعديل حتى عام 1939 حيث عُدلت في صورتها التي تستخدم حتى الآن. وقد كانت هذه الاختبارات نتيجة للبحوث الواسعة النطاق في ميدان سيكولوجية الموسيقى والتي أجريت بجامعة يوربا تحت إشراف كارل سي شور. وتتألف هذه البطارية في صورتها الحديثة من 6 اختبارات هي تمييز الأصوات، شدة الصوت، تذكر الإيقاعات، الزمن، نوعية الصوت، تذكر الألحان.

ويتكلف اختبار تمييز الأصوات من 50 زوجاً من النغمات، وعلى المفحوص أن يحدد بالنسبة لكل زوج ما إذا كانت النغمة الثانية أهد أم أعلظ من النغمة الأولى وفي اختبار شدة الصوت يطلب منه أن يحدد بالنسبة لكل زوج من 30 زوجاً من النغمات ما إذا كانت النغمة للثانية أشد أم أضعف من الأولى. ويطلب منه في اختبار تتكرر الإقاعات أن يحدد ما إذا كان النمطان الإقاعيان اللذان يتألف منهما كل سؤال (من 30 سؤالاً) متشابهين أو مختلفين. وفي اختبار الزمن يحدد للمفحوص ما إذا كانت النغمة للثانية أطول أم أقصر من النغمة الأولى (50 سؤالاً) وفي اختبار نوعية الصوت يقوم المفحوص بالحكم على النغمة الثانية بأنها متشابهة أو مختلفة عن النغمة من حيث نوعية الصوت (البناء الهارموني)، ويتكون هذا الاختبار من 50 سؤالاً. أما الاختبار الأخير، وهو تذكر الألحان، فيتكون من 30 زوجاً من النماذج اللحنية، وفي كل زوج توجد نغمة واحدة مختلفة في النموذج الثاني، وعلى المفحوص أن يحدد أي النغمات اختلفت ويحدد ترتيبها الرقمي.

وتصلح هذه الاختبارات للتطبيق على المفحوصين لبدء من الصف الرابع الابتدائي حتى أعلى المستويات التي تتطلب إعداداً موسيقياً رفيعاً للمراهقين والراشدين.

وقد ظهرت هذه الاختبارات في أصلها الأجنبي مسجلة على أسطوانات، وقد أعدتها باللغة العربية الدكتورة آمال أحمد مختار صادق مسجلة على شرائط، وأعدت تعليمات تطبيقها، كما صممت كراسة إجابة مختلفة عن الطبعة الأجنبية. وحسبت معاملات الثبات بطريقة التجزئة للنصفية وتصحيح سبيرمان براون على عينة من 163 طالباً وطالبة من حملة الثانوية العامة المتقنين

لالتحاق بالمعهد العالي للتربية الموسيقية وترواحت معاملات الثبات بين 58 واختبار الزمن 88 واختبار تذكر الأغنان. أما عن صدق هذه الاختبارات فقد حسبت الباحثة معاملات الارتباط بين الاختبارات الفرعية التي تتألف منها البطارية لعينات من طلاب دور المعلمين والمعلمات وطلاب المعاهد الموسيقية العليا وتلاميذ المدارس الإعدادية العامة، وتلاميذ المعهد القومي للموسيقى، كما حسبت معاملات الارتباط بين اختبارات سيشور واختبارات القدرات للموسيقى للأطفال (بنثلي) لعبة تلاميذ المرحلة الإعدادية العامة، وتلاميذ المعهد القومي للموسيقى، وكذلك معاملات ارتباطها باختبارات ونج للذكاء الموسيقي، وأوريجون للتذوق الموسيقي، وفارم للتكوين الموسيقي، وكامل للذكاء العام، واختبار الذكاء المصور للدكتور أحمد زكي صالح، وتتوافر في كراسة بالتعليمات بيانات كافية عن المعالم الميكومترية لهذه البطارية. أما عن معايير هذه البطارية فقد حسبت المنينيات الخاصة بكل اختبار فرعي ولكل عينة فرعية من عينة التتئين.

ويوجد اختبار آخر للقدرة الموسيقية ظهر عام 1963 أعده في إنجلترا أرنولد بنثلي خصيصاً للأطفال للمرحلة الابتدائية، (من 7-12 سنة) ويقسم 4 قدرات هي: تمييز الأصوات، تذكر النغمات، تحليل التناغمات، تذكر الإيقاعات. ويمكن تطبيق هذه الاختبارات على مستويات عمرية أعلى مع مراعاة معايير هذه الأعمار عند تفسير الدرجات.

ويتكون لاختبار تمييز الأصوات من 20 زوجاً من الأصوات ' للنقطة ' يستمع إليها المفحوص ثم يطلب منه أن يحدد ما إذا كان الصوت الثاني في كل زوج متشابهاً مع الصوت الأول أو مختلفاً عنه. وفي حالة الاختلاف على

المفحوص أن يحدد ما إذا كان أعلى (أي أحد) أو أسفل (أي أخفض) من الصوت الأول. أما اختبار تذكر النغمات فيكون من 10 أسئلة يتطلب كل منها مقارنة ثنائية بين جملتين لحظيتين قصيرتين، ويتكون كل منها من 5 نغمات، ويطلب من المفحوص أن يحدد ما إذا كانت الجملة الثانية متشابهة مع الأولى أو مختلفة عنها وإذا حدث اختلاف عليه أن يحدد ترتيب النغمة المختلفة. وفي اختبار تحليل التكاليف يطلب من المفحوص تحديد عدد النغمات التي يتكون منها كل تآلف (يتكون الاختبار من 20 تآلفاً). أما اختبار تذكر الإيقاعات فيتكون من 10 مفردات يتطلب كل منها مقارنة ثنائية بين جملتين إيقاعيتين، وتتكون كل واحدة من 4 وحدات إيقاعية أو ما يعادلها من النماذج الإيقاعية المختلفة، ويطلب من المفحوص تحديد ما إذا كانت الجملة الثانية في كل سؤال متشابهة مع الجملة الأولى أو مختلفة عنها، وفي حالة الاختلاف يكون عليه تحديد رقم للوحدة المختلفة.

وقد أعدت هذه الاختبارات بلقطة العربية للدكتورة آمال أحمد مختار صادق عام 1972، وظهرت اختبارات القدرات الموسيقية للأطفال مسجلة على شرائط (بدلاً من الاسطوانات في الأصل الأجنبي) وأعدت لها كراسة إجابة خاصة، كما فصلت تعليمات تطبيقها. وقد استخدم المؤلف الألفي في حساب ثبات هذه الاختبارات طريقة إعادة الاختبار، وبلغ معامل الارتباط 0.84. واستخدمت الباحثة المصرية هذه الاختبارات في دراستين عاملتين للقدرة الموسيقية إحداهما على عينة من تلاميذ المرحلة الإعدادية، والأخرى على عينة تناظرها في العمر الزمني من تلاميذ المعهد القومي للموسيقى (الكونسيرفطوري). عددها 74 تلميذاً.

ولحساب صدق الاختبارات حسبت معاملات الارتباط بين الاختبارات الفرعية، ومعاملات ارتباطها باختبارات سيشور التي أشرنا إليها، وبالاختبارات كاتل للنكاء العام. وكذلك حسبت المعايير المئينية لمؤنيتين من للتلاميذ: إحداهما مجموعة 9-12 سنة والأخرى مجموعة 12-15 سنة.

ويوجد اختبار آخر ظهر في إنجلترا عام 1941 ثم ظهرت طبعته الجديدة المعدلة عام 1958 ثم عام 1962 وهو اختبار ونج للنكاء الموسيقي.

ويصلح للتطبيق ابتداء من سن 8 سنوات، ويختلف عن اختبار سيشور في أنه لا يؤكد الجانب الذري التحليلي كما يتمثل في عناصر الإحساس البسيط المساعد في اختبارات سيشور، ولذلك فإنه يعتمد على محتوى موسيقي واقعي حيث تستخدم موسيقى البيلاو في كل اختبار من الاختبارات السبع التي تتكلف منها بطارية ونج وهي: تحليل التكتلات، تمييز الأصوات، ذاكرة الأصوات، التوافق الموسيقي (الهارموني)، الشدة، الإيقاع، الجمل الموسيقية. وتطلب الاختبارات الثلاث الأولى تمييزاً حسيّاً على مستوى أكثر تعقداً من اختبارات سيشور، أما في الاختبارات الأربع الأخرى فإن المفحوص يقارن للنوعية الجمالية في مقطوعتين موسيقيتين، ومعنى ذلك أن بطارية ونج تؤكد جانب اللتنوع الموسيقي. ويقترح ونج استخدام الاختبارات الثلاث الأولى كبطارية مختصرة مع الأطفال الصغار أو كأداة تصفية مبدئية.

اختبارات القدرة في الفنون البصرية

توجد في الوقت الحاضر طائفة من الاختبارات التي تقيس بعض جوانب القدرة ليميدان الفنون البصرية، ومنها اختبار Graves Design

Judgment الذي يهدف إلى قياس قدرة المفحوص على إدراك المبادئ الأساسية للعمل الجمالي والاستجابة لها وهي مبادئ الوحدة والمبادرة والتنوع والتوازن والاستمرار والتناغم والتناسب والإيقاع. وقد تحقق ذلك بمرض 90 ثنائية أو ثلاثية من الرسوم المجردة، في كل منها واحد من الرسوم يتفق مع المبادئ الثماني هذه بينما يختل أحد هذه المبادئ أو بعضها في الرسم الآخر أو الرسوم الأخرى. ويطلب من المفحوص أن يفاضل بين رسوم الثلاثية أو الثلاثية ويختار أحدها على أساس التفضيل الفني. ولم يتضمن هذا الاختبار نماذج من أعمال الفن، لأن الحكم على الصور واللوحات يتأثر بالخبرة السابقة وسمات الشخصية والانفعالات ومفهوم المفحوص عن الفن. ولذلك فإن الاختبار يؤكد محض الاختبار الجمالي، أو ما يسمى (الحساسية الجمالية). وبالإضافة إلى ذلك فإن الاختبار يقيس العمليات الإدراكية المعقدة وليس العناصر الحسية البسيطة التي تتدخل في الحكم الجمالي، وفي هذا يتشابه مع اختبارات ولح في الموسيقى، ومع المبادئ المتضمنة في قياس العامل في الذكاء.

والاختبار الثاني في هذا المجال هو اختبار ماير للحكم الفني. ويهدف إلى قياس الحكم الجمالي بطريقة (كالية). ويختلف عن اختبار جريغز في أنه يتألف من 100 زوج من اللوحات الفنية مرسومة بالأبيض والأسود، ويمثل أحد عنصرَيِ الثنائية صلاً فنياً ممتازاً، والعنصر الآخر أدخل عليه بعض التغيير والتعديل في أحد الجوانب الفنية الهامة تملأ أقل جودة من العنصر الأول. ويحدد للمفحوص أي الجوانب قد تغير (الشكل أو الزوايا مثلاً) إلى أنه لا يعلم أي اللوحتين هو اللوحة الأصلية، ويطلب من المفحوص أن يحدد تفضيله لأي منهما.

ويرى ماير أن هذا الاختبار يقيس الحكم الجمالي، إلا أنه في رأي أحد مؤلف هذا القياس يقيس (الحساسية الجمالية) أيضاً والفرق بينه وبين اختبار جريفز فرق في المحتوى وليس في العمليات الأساسية المتضمنة.

وإذا كان اختبار جريفز وماري ينتميان إلى ميدان للتذوق الفني فإن اختبري كنوير وهورن ينتميان إلى ميدان الابتكار الفني. ويسمى أولهما Knauber Arc Ability ويصلح للمراقبين والراشدين، ويتطلب من المفحوص ما يأتي:

- 1- أن يرسم من الذاكرة أشكالاً في حدود المكان المسموح به.
- 2- أن يرسم شخصيات لمطية (بلا تويل مثلاً).
- 3- ترتيب مؤلف معين في حدود المكان المسموح به.
- 4- ابتكار وتكملة رسوم معينة من عناصر معطاة.
- 5- تحديد الأخطاء في أعمال مرسومة (مثل المنظور الخاطئ).
- 6- إنتاج أعمال فنية تدل على الخيال والحق والتمثيل الرمزي.

أما اختبار هورن فقد أعد خصيصاً للاستخدام في أغراض قبول طلاب كليات الفنون الجميلة، ويقيس على وجه الخصوص الابتكارية في الفن مستخدماً في كل سؤال وصفاً بسيطاً مرناً أو مرشداً لكل رسم. ويتألف الاختبار من قسمين: الأول يتطلب من المفحوص أن يرسم 20 شيئاً مألوفاً (منزل، كتاب... إلخ) في وقت قصير، كما يتطلب منه إنتاج رسوم تجريدية بسيطة مبتكراً بمجموعة من المثلثات والمربعات وغيرها. أما القسم الثاني فيُطلب مستوى أعلى من الابتكارية وفيه يعطى للمفحوص 20 بطاقة مستطيلة على كل منها

خطوط تفيد كبدائية لصورة ينتجها المفحوص. ويحكم على الرسوم على أساس الخيال الابتكاري من ناحية والخصائص الفنية من ناحية أخرى.

ومن الاختبارات التي تقيس القدرات في الفنون البصرية اختبارات تكميل الأشكال وتكميل الصور والتقدير الجمالي التي أعدها الدكتور محمد إسماعيل لقياس عاملي الطلاقة والتقدير الجمالي. وقيس اختبار تكميل الأشكال عامل الطلاقة وفيه يطلب من المفحوص أن يضيف إلى خطوط مرسومة بضعة خطوط أخرى بحيث تجعل منه صورة لشيء مألوف. وقيس اختبار تكميل الصور عامل الطلاقة أيضاً وفيه يطلب من المفحوص أن يكتب في مكان معين من صورة عامل مرسومة أكبر عدد من الأشياء التي يمكن رسمها. أما اختبار التقدير الجمالي فهو تقتين مصري لاختبار ماير، الذي أشرنا إليه. وقد حسب صدق الاختبارات بإيجاد معاملات الارتباط بين كل منها وبين المحك (وهو تقدير المدرسين للقدرة الفنية) فكانت النتائج كالآتي: 0.44 لاختبار تكميل الأشكال، 0.34 لاختبار تكميل الصور، 0.44 لاختبار التقدير الجمالي (ماير).

اختبارات التفكير الابتكاري

من أهم خصائص التقويم النفسي المعاصر زيادة الاهتمام بقياس الابتكارية مسايرة للنشاط المتزايد في مختلف موانع علم النفس لدراسة طبيعة النشاط الابتكاري والعوامل المؤثرة فيه. ولا يتسع المقام للإفاضة في هذه المسائل، ويمكن الرجوع إلى بعض المصادر المتاحة في اللغة العربية ونستطيع القول بإيجاز أن الاهتمام بالتفكير الابتكاري يعبر عن حاجة المجتمعات المعاصرة لزيادة وتنمية ثروتها البشرية من الطعام والمهندسين والقادة الإداريين والفنانين. ويرى كثير من العلماء للمهنيين بهذا الميدان أن الابتكار عملية

أساسية في هذه الأحوال جميعاً، فهو كما يقول تايلر (يختلف في العمق والمجل وليس في النوع) ومن رأيه عدم للتمييز بين الابتكار العلمي والابتكار الفني لأن الابتكار (يتضمن نظرة للمشكلات أكثر أساسية من التدريب المهني).

ويعود الفضل إلى العالم الأمريكي المعاصر جيلفورد في زيادة حركة قياس التفكير الابتكاري، فقد أثرت نظريته عن تنظيم العقل في معظم بحوث الابتكار، وفيها يقترح تصنيف القدرات العقلية إلى فئتين كبيرتين هي قدرات للتفكير، وقدرات للذاكرة. وتتقسم عوامل التفكير إلى ثلاثة أنواع هي:

1- قدرات التفكير المعرفي (بمعنى الاكتشاف).

2- وقدرات التفكير الإنتاجي.

3- قدرات التفكير التقويمي.

أ- قدرات التفكير الإنتاجي التقاربي.

ب- قدرات التفكير الإنتاجي التباعدي.

والإنتاج التقاربي (أو الاختزالي) يتضمن تضيق الاحتمالات عند إنتاج

إجابة واحدة محتملة للمشكلة، مثل: طويل إلى قصير مثل مرتفع إلى ... ؟ وبالطبع فإن أغلب اختبارات الذكاء العام تقوس هذا النمط من التفكير، بل إلى معظم الاختبارات العقلية التي أشرنا إليها في الفصول السابقة من هذا الكتاب تنتمي إلى فئة الإنتاج التقاربي. أما التفكير التباعدي فيتطلب إنتاج أكبر عدد ممكن من الإجابات. مثل: ما هي الاستخدامات التي يمكن أن تفكر فيها لقلب الطوب ؟.

ويرى جيلفورد أن الإنتاج التباعدي عامل هام في التفكير الابتكاري. واستخدم ووضع اختبارات عديدة لقياس هذه القدرة في مستويات مختلفة وفي محتويات متباينة. ومن رأيه أن بعض عوامل الذاكرة والتفكير المعرفي والتفكير التقويمي والتفكير الإنتاجي التقاربي ترتبط بالسمل الابتكاري.

ويعتقد على وجه الخصوص أن الموهبة الابتكارية يمكن قياسها بالاختبارات التي تقيس قدرات الأصالة والطلاقة والمرونة من مختلف الأنواع، وقدرات إعادة التحديد والتضمنين والتفكير التقويمي أيضاً.

وقد أثرت نظرية جيلفورد في ظهور بطاريتين رئيسيتين لقياس التفكير الابتكاري إحداهما أعدّها جيلفورد وزملاؤه بجامعة كاليفورنيا. والأخرى أعدّها تورنس بجامعة مينيسوتا.

اختبارات جيلفورد للتفكير الابتكاري

تتألف بطارية اختبارات التفكير الابتكاري لجيلفورد على النحو الآتي:

1- اختبارات جيلفورد للطلاقة وتتألف من 4 اختبارات يطلب فيها من

المفحوص أن يكتب كلمات بأسرع ما يمكن تحت الشروط الآتية:

أ- طلاقة للكلمات (الإنتاج التباعدي لوحدات الرموز)، يطلب فيه من

المفحوص أن يكتب كلمات ذات خصائص معينة، كأن تتضمن

حرف (ب) مثلاً.

ب- طلاقة الأفكار (الإنتاج التباعدي لوحدات المعاني): يطلب فيه من

المفحوص أن يذكر أسماء الأشياء التي تنتمي إلى فئة معينة، مثل

السوائل التي تحترق.

ج- طلاقة التداعي (الإنتاج التباعدي للعلاقات بين المعاني): يطلب فيه من المفحوص إعطاء كلمات تتشابه في المعنى مع كلمة أخرى.

د- الطلاقة التعبيرية (الإنتاج التباعدي لمنظومات الرموز): يتطلب فيه من المفحوص تكملة جمل، كل منها يتألف من 4 كلمات، بحيث تبدأ كل كلمة بحرف معين.

2- اختبار الاستخدامات البديلة (الإنتاج التباعدي لصفات المعاني): يطلب فيه من المفحوص أن يذكر الاستخدامات المحتملة لشيء معين غير استخدامه المألوف (قالب الطوب مثلاً يستخدم في البناء استخداماً مألوفاً).

3- اختبار المترتبات (الإنتاج التباعدي لوحدات المعاني، والإنتاج التباعدي لتحريكات المعاني) وفيه يطلب من المفحوص أن يذكر للنتائج المختلفة المترتبة عن حادثة افتراضية (مثل: ماذا يحدث لو أن الناس توقفوا عن النوم؟). ويطلب للمفحوص في هذا الاختبار درجتان: إحداهما للاستجابات الواضحة وتقيس طلاقة الأفكار أو الإنتاج التباعدي لوحدات المعاني، والأخرى للاستجابات البعيدة وتقيس الأصالة، والإنتاج التباعدي لتحريكات المعاني.

4- الأعمال المحتملة (الإنتاج التباعدي لتضمينات المعاني)، ويطلب فيه من المفحوص أن يذكر الأعمال المحتملة التي يمكن أن يرمز لها بشعار (مصباح كهربائي مثلاً).

5- عمل الأشياء (الإنتاج التباعدي لمنظومات الأشكال): ويطلب فيه من المفحوص أن يرسم أشياء محددة باستخدام مجموعة معطاة من الأشكال (كالدوائر أو المثلثات). وعلى المفحوص ألا يستخدم إلا الأشكال المعطاة فقط.

- 6- الاستكشافات (الإنتاج التباعدي لوحدة الأشكال): ويشبه اختبار عمل الأشياء فيما عدا أن كل صفحة في الاختبار تشتمل على نفس الشكل (دوائر مثلاً فقط) ويطلب من المفحوص أن يرسم أكبر عدد من الاستكشافات بالإضافة تفاصيل إلى كل شكل من الأشكال المعطاة.
- 7- مشكلات عيدان الثقاب (الإنتاج التباعدي لتحويلات الأشكال): وفيه يطلب من المفحوص أن يستبعد عدداً معيناً من عيدان الثقاب (المرسومة على الورق في صورة خطوط مستقيمة) بحيث ينتج عن ذلك عدداً من المربعات أو المثلثات.
- 8- اختبار الزخرفة (الإنتاج التباعدي لتضمينات الأشكال): وفيه يطلب من المفحوص أن يزخرف صوراً لأشياء ملوقة بأكثر قدر من الرسوم المختلفة.

ويلاحظ أن الاختبارات السبع الأولى من هذه القائمة تتطلب استجابات لفظية، بينما تستخدم الاختبارات الأربع الأخيرة. محتوى من الأشكال أو الصور. وهذه الاختبارات الأحد عشر تمثل عينة صغيرة من مجموعات الاختبارات التي أنتجتها بحوث جيلفورد وزملائه وتلاميذه بجامعة جنوب كاليفورنيا لأكثر من عشرين عاماً. وقد أعدت معايير مبدئية لهذه الاختبارات في صورة مبدئيات ومعايير جيمية C-sore (وهو معيار يلخص المستويات للتأقبة إلى 11 مستوى وهو في ذلك يشبه المعيار للتساعي ويختلف عنه في عدد المستويات). وفي معظم الأحوال اشتقت هذه المعايير لعينات من الراشدين أو طلاب من سن 15 + أو كليهما. وتصلح هذه الاختبارات بوجه عام لمستوى المرحلة الثانوية وما بعدها، رغم أن جيلفورد يذكر أن بعض اختباره تصلح لمستوى المرحلة الابتدائية (ابتداء من الصف الأول).

والمشكلة الجوهرية في اختبارات التفكير الابتكاري هي مشكلة التصحيح فهي تتطلب من المصحح تدريماً كلفياً على اتباع تعليمات التصحيح ودراسة الأمثلة التي تتضمنها كراسات التعليمات. ورغم ذلك فإن عملية تصحيح هذه الاختبارات تستغرق وقتاً كبيراً.

وحسب جيلفورد ثبات هذه الاختبارات بطريقة التجزئة النصفية، ويتراوح معاملات الثبات بين 0.60، 0.80 أما عن الصدق فقد اعتمد اعتماداً كبيراً على صدق التكوين الفرضي كما يتمثل في اتفاق الاختبارات مع نموذج العقل الذي يقترحه. ولا زالت هذه الاختبارات في حاجة إلى أنواع الصدق الأخرى وخاصة الصدق المرتبط بالمحتكات.

وقد قام أحد الدكاترة بإعداد بعض اختبارات جيلفورد باللغة العربية، فأعد اختبارين لقياس الطلاقة اللفظية (طلاقة للكلمات) هما اختبار الطلاقة اللفظية (1)، ويتطلب من المفحوص إنتاج كلمات تبدأ بحرف معين، واختبار الطلاقة اللفظية (2) ويتطلب من المفحوص أن ينتج كلمات تنتهي بحرف معين. كما أعد (3) اختبارات أخرى لقياس بعض عوامل التفكير الابتكاري الأخرى وهي: اختبار الطلاقة الفكرية (لقياس عامل طلاقة الأفكار)، واختبار الاستعمالات (لقياس عامل المرونة للتفكير) واختبار المترتبات (لقياس عامل الأصالة). وقد حسب ثبات هذه الاختبارات من أداء عينة عددها (120) تلميذاً بالمدارس الثانوية العامة بطريقة إعادة الاختبار لاختباري الطلاقة اللفظية فبلغ معاملات الثبات 0.62، 0.66 وباستخدام في حساب ثبات اختبارات لطلاقة الفكرية والاستعمالات والمترتبات طريقة التجزئة النصفية وبلغت معاملات الثبات 0.75، 0.69، 0.80 على التوالي.

وفي تقدير الصديق حسب الباحث معاملات ارتباط الاختبارات بالتحصيل المدرسي، وتراوحت هذه المعاملات بين 0.27 (الاستعمالات)، 0.58 (المترتبات)، ولم تتضمن كراسة بالتعليمات بيانات عن معايير الاختبارات.

اختبارات تورنس للتفكير الابتكاري

أشرنا إلى أن اختبارات جيلفورد هي محصلة بحوثه العملية في ميدان التفكير الإنساني وطبيعة الذكاء، أما اختبارات تورنس للتفكير الابتكاري التي نتناولها الآن فقد ظهرت لمواجهة للضرورات التربوية والعملية كجزء من برنامج طويل من البحوث تهتم بالخبرات التعليمية التي تساعد على تنمية الابتكار، وبعض اختبارات تورنس هي الواقع تعديل للأساليب التي استخدمها جيلفورد بالإضافة إلى أن الدرجات المشتقة من البطاريات الكاملة تعتمد على العوامل التي توصلت إليها بحوث جيلفورد وهي الطلاقة والمرونة والأصالة وإعطاء التفاصيل. وفي هذا الصدد يذكر تورنس أنه لم يسع إلى وضع اختبارات (نقية عاملياً) أو (بسيطة التكوين)، وإنما إلى التوصل إلى اختبارات تتضمن مولف ماثلة لما تتطلبه العملية الابتكارية في الظروف الطبيعية المعقدة. ولذلك فإن كل اختبار قد يصحح في ضوء عاملين، أو ثلاث عوامل، أو العوامل الأربع جميعاً.

وتألف اختبارات تورنس للتفكير الابتكاري من (10) اختبارات، مصنفة إلى بطارتين: بطارية لفظية وبطارية مصورة. وتسمى البطارية الأولى التفكير الابتكاري بالكلمات، والثانية: التفكير الابتكاري بالصور. وحتى يمكن التخفيف من التهديد الذي تتضمنه كلمة (الختبار) استخدمت في كل الأحوال كلمة (نشاط)، كما تؤكد التعليمات على (اللعب)، وتوصف هذه الاختبارات بأنها

تصلح لمختلف المستويات ابتداء من مرحلة الحضنة حتى مستوى الدراسات العليا، بشرط أن تطبق لردياً وشفوياً على المستويات الأدنى من الصف الرابع الابتدائي. وتوجد صورتان متكاملتان لكل بطارية من البطاريتين.

وفي التفكير الابتكاري بالكلمات: الصورة (أ) والصورة (ب)، تتألف الأنشطة الثلاث الأولى (لأسل وخمن) من صورة غير عادية يستجيب لها المفحوص بكتابة:

- أ- جميع الأسئلة التي يحتاج إلى توجيهها لكي يعرف ما يحدث في الصورة.
- ب- الأسباب المحتملة لما يحدث في الصورة.
- ج- النتائج المترتبة على ما حدث في الصورة.

أما النشاط الرابع فيتعلق بطرق تحسين لعبة من لعب الأطفال بحيث تصبح مصدراً لمزيد من السرور والفرح لمن يلعب بها. ويتعلق النشاط الخامس بذكر الاستخدامات غير العادية لشيء ما؛ المؤلف: (علب الكرتون) في الصورة (أ)، (علب الصفيح) في الصورة (ب). ويتعلق النشاط السادس بالأسئلة غير المعتادة التي يمكن توجيهها عن الشيء المتضمن في النشاط الخامس. أما النشاط السابع فقد أعد على نسق اختبار المترتبة على جيلفورد، ويتطلب من المفحوص أن يذكر كل ما يمكن أن يحدث إذا نشأ موقف غير ممكن الحدوث. وتطلى البطارية درجات كلية في العوامل الثلاث: المرونة والأصالة.

ويتألف اختبار (التفكير الابتكاري: الصورة أ، والصورة ب من ثلاث أنشطة:

1- تكوين الصورة وفيه يعطى للمفحوص ورقة ملونة ذات شكل منحنى وعليه أن يضعها على صفحة بيضاء في الموضع الذي يشاء ويستخدمها كبداية لرسم صورة غير عادية (والورقة الملونة المستخدمة في الصورة (أ) زرقاء اللون، والمستخدم في الصورة (ب) صفراء اللون في الطبعة العربية).

2- تكملة للصورة: ويتألف من (10) خطوط مختلفة يستخدم كلاً منها في رسم صورة منفصلة، وهذا الاختبار يستخدم في نفس الأسلوب المستخدم في اختبار هورن للاستعداد الفنى الذي أشرنا إليه، مع فرق هام هو أن تعليمات اختبار تورنس تؤكد لإنتاج الأفكار غير المعتادة، كما أن التصحيح يعتمد على جوانب الابتكارية وليس الخصائص الجمالية والفنية.

3- يعطى النشاط الثالث للمفحوص قائمة من الخطوط المتوازية (للمصورة (أ) أو الدوائر (للمصورة ب))، وعليه أن ينتج أكبر عدد من الصور والرسوم.

ويحصل المفحوص في البطارية المصورة على (4) درجات هي:
الطلاقة المرونة والأصالة وإعداد التفاصيل.

وقد ظهرت هذه الاختبارات في عام 1971 على النحو الآتي:

1- التفكير الابتكاري بالكلمات: الصورة أ.

2- التفكير الابتكاري بالكلمات: الصورة ب.

3- التفكير الابتكاري بالصورة: أ.

4- التفكير الابتكاري بالصورة: ب.

- 5- كراسة تعليمات اختبارات تورنس للتفكير الابتكاري.
- 6- استمارة تصحيح (1) لاختبارات التفكير الابتكاري بالكلمات (الصورة أ، والصورة ب).
- 7- استمارة تصحيح (2) لاختبارات التفكير الابتكاري بالصور (الصورة أ، والصورة ب).
- 8- استمارة تقدير المعلمين للابتكار.
- 9- استمارة تقدير التلاميذ للابتكار.

وقد طبقت هذه الاختبارات على عينة كبيرة من تلاميذ المرحلتين الإعدادية والثانوية في العام الدراسي (1972 - 1973)، ويقوم الباحثان في الوقت الحاضر بإعداد معايير التصحيح، وتحديد المعالم الميكومترية تمهيداً لنشرها في كراسة منفصلة تمهيداً لنشرها، وقد صدرت الكراسة الأولى عام 1973 وتتضمن تقنين الصورة أ من اختبارات الأشكال أو للصور، فطبق الاختبار على 313 تلميذاً وتلميذة من مدرستين إعداديتين بالقاهرة تتراوح أعمارهم بين 12، 16 سنة، وحسب الصدق استخدمت استمارة تقدير المعلمين للابتكار حيث طلب من مدرس الفرق التي أجريت عليها الاختبارات أن يختار كل منهم (5) تلاميذ يقعون في الفئة العليا، (5) تلاميذ يقعون في الفئة الدنيا من كل بعد من أبعاد الابتكار المقاسة: الطلاقة، المرونة، الأصالة، للتفاصيل. واختير كل تلميذ يتفق عليه مدرسان على الأقل. وحسبت دلالة الفروق بين متوسطات مجموعات للتلاميذ الأعلى والأدنى في مختلف أبعاد الاختبار باستخدام مقياس (ت) فوجنت للفروق في معظم الأحوال (باستثناء للتفاصيل) دلالة وفي الاتجاه الصحيح إلا أن المستوى للدلالة كان عالياً (0.01) بالنسبة إلى بعد المرونة، وكان منخفضاً نسبياً (0.10) بالنسبة إلى بعد الطلاقة والأصالة.

ويستخدم في حساب الثبات الطرق الآتية:

- 1- حساب معاملات الارتباط بين متغيرات الاختبار (الطلاقة، المرونة، الأصالة، التفاصيل، الدرجة الكلية) فوجنت المعاملات في كل الأحوال ذات دلالة إحصائية عند مستوى (0.01).
- 2- حساب معاملات الارتباط بين مصححين قلماً بتصحيح (38) اختباراً كل على حدة، وكانت معاملات ثبات التصحيح (0.99) للطلاقة (0.99) للمرونة ولأصالة، (0.97) للتفاصيل، (0.99) للدرجة الكلية ويرجع ارتفاع المعاملات على هذا النحو إلى التزم المصححين بقواعد مقننة في التصحيح، كما لجأ الباحثان في تدريب المصححين إلى طريقة تورنس، وتأكدت بالفعل جدواها.
- 3- حساب الثبات بطريقة إعادة الاختبار وبلغت المعاملات 0.54 للطلاقة، 0.36 للمرونة، 0.53 للأصالة، 0.40 للتفاصيل، 0.50 للدرجة الكلية في الاختبار.

وقد أعد الباحثان معايير الاختبار في صورة درجات ثائية، كما أعد للنماذج التجريبية لتصحيح الصورة (أ) من اختبارات الأشكال، وهي نماذج مشتقة من استجابات عينة التقنين المصرية ذاتها وليست منقولة عن كراسات التعليمات الأجنبية.

قياس الشخصية عن طريق الأحكام

قياس الشخصية

قياس الشخصية عن طريق الأحكام والملاحظة المنظمة:

رأبنا في قياس الشخصية عن طريق التقدير الذاتي أن الشخص يختار ويحطي تقديرات يرى أنها تعبر عن نفسه، أما عند قياس الشخصية عن طريق الأحكام والملاحظات المنظمة فإن الذي يقوم بتقدير السلوك أو الأداء المميز للفرد أو للمفحوص شخص أو أشخاص آخرون، ويمكن أن نقسم هذا النوع من قياس الشخصية إلى قسمين:

- أ- مقاييس التقدير وتقديرات للقياس الاجتماعي.
- ب- الملاحظة المنظمة لعينات السلوك.

وسوف نتناول كلا من هذين القسمين بشيء من التفصيل:

أ- مقاييس التقدير وتقديرات للقياس الاجتماعي:

يمكن أن نصنف مقاييس التقدير وتقديرات القياس الاجتماعي إلى نوعين على أساس من يقومون بهما، للنوع الأول يقوم به الرؤساء أو الكبار، بينما يقوم الأفراد أو الزملاء بالتقدير والحكم في النوع الثاني.

1- تقديرات الرؤساء أو الكبار:

نحن نلاحظ أن الرؤساء أو المدرسين مثلاً يقومون باستمرار بتقييم نوع من الأوصاف لمرووسهم أو لمن هم في كفالتهم، إلا أننا لا نستطيع أن

ن هذه الأوصاف لأنها تختلف في أسلوبها، لهذا استعملت مقاييس الأ
ى بتوفر قدر معقول من التشابه بين التقديرات.

ويتكون مقياس التقدير من قائمة من السمات يجرى عليها التقدير.
ون للمقياس من قائمة بسيطة من الصفات يتم الاختيار بينها، أو قد و
باماً متصلاً يتضمن أسماءً وصفية متعددة للاختيار منها. ويوضح ال
لي، هذا النوع من مقاييس التقدير المتصلة:

وقائع مؤيدة	<ul style="list-style-type: none"> • يسمى إليه الآخرون. • يحبه الآخرون. • يحتمله الآخرون. • يتجنبه الآخرون. • لا مجال لملاحظته. • يسمى إلى إضافة أعمال جديدة إلى نفسه. • يكمل للعمل الإضافي المقترح. 	<p>1- كيف تتأثر أنت والآخرون بمظهره وسلوكه ؟</p>
وقائع مؤيدة	<ul style="list-style-type: none"> • يعمل ولجبهه العادية تلقائياً. • يحتاج إلى تحريك أحياناً. • يحتاج إلى تحريك كثير ليعمل ولجبهته العادية. • لا مجال لملاحظته. 	<p>2- هل يحتاج إلى تحريك أو أنه يعمل من تلقاء نفسه ؟</p>

نلاحظ أن مقاييس التقدير عن طريق الرؤساء تكمن فيها مصادر متعددة للخطأ منها مثلاً ما يسميه كرونباك (خطأ للكرم) ويقصد به ميل من يقوم بالتقدير إلى إعطاء تقديرات طيبة مؤيدة، فإذا مثل مدرس مثلاً أن يقدر تلميذاً ما فيما يختص بالميل إلى التعاون، فإنه سوف يقدر كل التلاميذ الذين لا يثيرون للشغب تقديراً عالياً جداً، في العكس طرف المقياس، بينما يهبط بتقدير لقلّة من التلاميذ مثيри للشغب، ويكون الارتفاع في تقدير الغالبية رد فعل تمييزهم عن مثيري الشغب. ومن أسباب خطأ للكرم أيضاً أن الرئيس المقدر قد يحس أن مما يمس قدرته الرئاسية أن يعترف أن مروضيه ليسوا أكفاء، أو أن الرئيس قد يجد من الأسر عليه أن يحكم أحكاماً طيبة من مروضيه إجمالاً يدل أن يتملى ويطلق النظر ليميز بينهم تمييزاً دقيقاً.

ومن مصادر الخطأ أيضاً في هذا النوع من التقدير الغموض والإبهام، فقد يفسر المقدر سؤالاً أو مصطلحاً تفسيراً مختلفاً، فقد يفسر شخص (القيادة)، مثلاً على أنها تركيز السلطة في يد القائد، والافراد باتخاذ القرارات، فعندما يقدر هذا الشخص مفحوصاً تتوافر فيه هذه الصفات فإنه يقدره تقديراً عالياً ومخالفًا لتقدير شخص آخر للمفحوص ذاته، لأن المقدر الأخير يفهم القيادة على أنها تشجيع المرؤسين ومساعدة على اتخاذ قرارات جماعية.

ويعتبر أسلوب الاستجابة عند المقدر مصدراً آخر من مصادر الخطأ في مقاييس التقدير بواسطة الرؤساء أو الكبار، فمثلاً نجد أن مقدار ما يندر أن يستعمل أطراف المقياس في وصف المفحوصين، بينما يميل شخص آخر إلى الإسراف في استعمال هذه الأطراف، أي يميل إلى تصنيف المفحوصين إلى

لقسام متميزة جداً. ومن ثم نجد أن مثل هذين المقدارين يختلفان علماً بقدران
شخصاً واحداً.

ومما يؤثر في تقديرات الرؤساء أيضاً أن معلوماتهم عن المفحوصين قد
تكون محددة، فقد يختلف تقدير مدرس للتربية الفنية ومدرس اللغة العربية في
تقدير ردود فعل تلميذ معين للهبوط ويرجع الاختلاف إلى أن كلا منهما يعرف
التلميذ ويلاحظه في عدد محدود من المواقف، وعلى هذا تكون معرفته به
محدودة وليست شاملة.

ومن مصادر الخطأ في هذا النوع من التقدير (أثر الهالة) أو (وقع
الحالة). ويحدث هذا الأثر أو الوقع عندما يتأثر المقدر تأثيراً قوياً بالاتطباع العام
عن المفحوص، فيتكون عنده رأي عام أو تقدير عام عن قيمة الشخص يؤثر
كثيراً في تقديره لسمات هذا الشخص للخاصة. وهذا النوع من الخطأ يجعل نمط
السمات في لف رد شامخاً عند المقدر.

إلا أنه من الممكن تلافي كثير من مصادر الخطأ هذه إذا راعينا النقاط التالية:

1- الاختيار السليم لمن يقون بالتقدير، من حيث قدرتهم ودرجة اندماجهم
وتأثرهم شخصياً بتقديرهم للمفحوصين، فقد سبقت الإشارة إلى أن
المقدر قد يميل في تقديره إلى التحيز لصالح مروضيه لأن في هذا
تأييد لكفاءته.

2- كلما زادت معرفة المقدر بالمفحوصين كانت تقديراته أكثر صدقاً.

3- إذا استعملنا مقدراً بصله أكثر من مرة فمن المفيد أن تحتفظ بتقديراته لمقارنتها والتعرف على أسلوب استجابته، أو ما يمكن أن نسميه خطأه الثابت.

4- من الممكن زيادة ثبات مقاييس التقدير عن طريق الاعتماد على أكثر من مقدر واحد، على أن يكونوا على معرفة كافية بالمفحوصين.

5- مما يزيد من قيمة مقاييس التقدير أن تعد بعناية ودقة. وبصفة عامة نجد أن المقياس ذا الخمس نقط يعطي تمييزاً أكثر دقة من المقياس أو القائمة التي تعتمد على إجابات نعم أولاً. كما أن مقياس الخمس نقط ينه إلى أنواع مقايينة من الاختلافات والانحرافات. وكذلك يساعد الاختيار الإجباري الذي نلكنما عنه في بعض اختبارات التقرير لذاتي، على زيادة قيمة مقياس التقدير. والهدف من الاختيار الإجباري هو فصل عمل المقدر فيما يتصل بوصف ما يعمله المفحوص عن تقويم هذا العمل، أي أن المقدر يكون مسئولاً عن وصف المفحوص، بينما يقوم المسئول عن اتخاذ القرار بتقويم هذا الوصف. وهذا النوع من مقاييس التقدير للقائم على الاختيار الإجباري أكثر ملاءمة لقرارات المؤسسات فيما يتصل بالانتقاء والتصنيف، ولكنها لا تناسب كثيراً صلية للتوجيه التربوي أو الوصف للتصليي للفرد.

لما عن صدق مقاييس تقدير الرؤساء فتقدير الدراسات التي أجريت عنها إلى ضرورة أخذ هذه المقاييس بحذر فيما يتصل بالحكم على الأفراد من جهة، وباستعمالها للتحقق من صدق بعض مقاييس واختبارات الشخصية من جهة أخرى. ومن أمثلة هذا أن طريقة تقدير الكبار استعملت في دراسة صدق مقياس المسئولية الاجتماعية، للصورة (ت)، الذي تناوله الفصل السابق، حيث

قام المدرسون بتقدير تلاميذ العينة على مقياس خالص بالمدرسين وكان معامل الارتباط بين تقدير المدرسين ودرجات التلاميذ على المقياس 0.46. وقد اعتبر هذا المعامل مقبولاً (وخاصة إذا أخذنا في الاعتبار أن تقدير المدرس للتلميذ في أي ناحية من نواحي السلوك تتدخل فيه عوامل كثيرة، منها على سبيل المثال، تحصيل التلميذ الدراسي، ومستواه الاقتصادي الاجتماعي، ودرجة جاذبية التلميذ الشخصية، وتفضيل المدرس وتحيزه الشخصي وهكذا).

وقد أثبتت دراسة قام بها بيترز وآخرون أن ارتباط تقديرات الرؤساء لمدى معرفة العاملين درابتهم بأعمالهم مع هذه المعرفة والدراية كما قيست باختبار خاص، كان الارتباط بينهما 0.35 بينما وصل هذا الارتباط في بعض الأقسام 0.55، كما أظهرت دراسات أخرى ارتباطاً لتقديرات بطول المدة التي يعرف فيها المقرر المفحوصين، وبمقدار حبه لهم.

ولكن مقياس تقدير الرؤساء والكبار، على الرغم من حدودها وتعتبر مصدراً غنياً بالمعلومات التي لا يمكن الحصول عليها إلا عن طريقها.

2- تقدير الأقران:

هي التي تعطي معلومات تعو في قيمتها على تقديرات الكبار أو الرؤساء، وحتى عندما تتوفر تقديرات الرؤساء فإن تقديرات الأقران والزملاء تعطي بيانات عن جلاب مختلف من الشخصية. فإذا كان رئيس واحد هو الذي يقوم بالتقدير فإن قيام مجموعة من الأفراد أو الزملاء بالتقدير سوف يعطي فرصة لعدد كبير من الأحكام مما يترتب عليه أن متوسط التقديرات بالنسبة إلى سمة معينة من سمات الشخصية سوف تكون أكثر ثباتاً من التقدير الذي يقوم به

رئيس واحد. ومن الممكن أن ننظر إلى تقديرات الأفراد على أنها تعبير موضوعي عن سمعة الفرد، فالسمعة تقوم إلى حد ما على السلوك الفعلي، إلا أن نمط العلاقات الاجتماعية في داخل الجماعة لا شك يؤدي إلى وجود نوع من التحيز.

ومن الأساليب المستعملة في تقدير الأفراد ما يسمى أسلوب تعيين الأسماء أو الترشيح، ويستعمل هذا الأسلوب كثيراً في القياس الاجتماعي وفي هذا الأسلوب يسأل كل عضو من أعضاء الجماعة أن يسمي عدداً محدداً من الأشخاص الذين يتميزون بشكل واضح في ناحية من النواحي، مثل القيادة، كما يسمى أيضاً الأشخاص الذين تتقصر هذه الخاصة. وباللمبة إلى الأطفال فإن عملية للتسمية قد تتخذ صورة لعبة للتخمين، أو اختبار (خمن من) الذي يصف الأكوام المختلفة التي يمكن أن يلعبها الأطفال، ويسمى كل عضو في الجماعة أسماء الأطفال الذين يعتقدون أن كل وصف يلائمهم. وترسم لكل طفل صفحة نفسية عن طريق عدد المرات التي أشير إليه فيها لكل وصف.

وإذا كان من الممكن دراسة خصائص الفرد بطريقة خمن من، فإن الرسم الاجتماعي (السوسيوجرام) يعطي فهماً واستبصاراً أعمق في تمييز للتجمعات للدخلية في الجماعة، وبيان البناء الهرمي للقيادة والمنعزلين، أي هذا الرسم الاجتماعي يعتبر طريقة لدراسة البنية الاجتماعية في الجماعة.

والذي أنشأ الرسم الاجتماعي وطوره مارينو، إلا أن الأسلوب قد عدل بعدة طرق قللت من فاعليته في سبيل السهولة واليسر في الاستعمال. وأفضل طريقة في استعماله هي أن يسأل أعضاء الجماعة أن يختاروا رفقاء لهم في نشاط معين. ولا بد أن تكون التعليمات متصلة بالنشطة حقيقة للجماعة، كما يجب

أن تكون الاختبارات حقيقية ويجب أن يعلم أعضاء الجماعة أن للمعلومات التي يعطونها سوف تصان سرّيتها، وأن البيانات المعطاة سوف تستخدم فيما أخبر أعضاء الجماعة أنها سوف تستخدم فيه، مثلاً لاختيار لجان، أو جماعات عمل فرعية، أو تنظيم للجلوس في فصل دراسي وهكذا.

ويستند للرسم الاجتماعي الذي نحصل عليه على السؤال الذي توجهه، فإن نمط الرسم الذي ينتج عن سؤال مجموعة من التلاميذ عن اختياراتهم لأقرانهم الذين يحبون أن يستذكروا معهم سوف يختلف عن نمط الرسم الناتج عن سؤال عن اختيارات من يحبون اصطحابهم في رحلة.

ويشمل مصطلح التقدير بالقياس الاجتماعي كل طرق تمييز العلاقات الاجتماعية بين أعضاء الجماعة. وليس هناك فصل قاطع بين تقدير الأفراد الوصفي، وبين التقدير بالقياس الاجتماعي، إلا أننا نجد أن النوع الأخير بصفة محدّدة بأسئلة عن يحبهم المقدر أكثر أو يفضل العمل معهم، ولهذا فهي تتصل باستجابة المقدر كما تتصل بشخصيته.

ولتقديرات الأقران استعمالات متعددة منها أن قائد الجماعة يستعملها في التعرف على الأشخاص الذين يحتاجون إلى عناية خاصة، أو يستعملها في إعادة تنظيم الجماعة حتى تؤدي وظيفتها بشكل أفضل. كما أن هذه التقديرات يمكن أن تتخذ أساساً للانتقاء والتصنيف، وهناك من يعتبرها أنقى مقياس للقيادة، أما بالنسبة إلى المرشد النفسي، فإن تقديرات الأقران توضح خصائص الشخص التي تعرق قبول الجماعة له، وخاصة عندما ينقص للشخص الاستبصار الكافي في رأي الآخرين فيه.

ويذهب كثير من الباحثين إلى أن تقديرات الأقران أثبتت أنها إحدى وسائل التقدير التي يمكن أن يوثق بها إلى درجة بعيدة لعدة أسباب هي:

1- أن عدد المقدرين كبير، وقد أشرنا منذ قليل إلى أهمية هذه في زيادة ثبات التقديرات بصفة عامة.

2- أن أقران الشخص عادة ما يكونون أكثر قدرة من غيرهم على ملاحظة السلوك المميز له، وهم من ثم أفضل في الحكم على سمات معينة تتصل بالعلاقات الاجتماعية من المدرسين أو المشرفين أو أي ملاحظين خارجين.

3- أن آراء أعضاء الجماعة الآخرين، صولياً كانت لم خطأ، تؤثر في سلوكهم وتصرفاتهم وبالتالي فإن هذه الآراء كتحد جزئياً طبيعة تفاعلات الشخص مع الجماعة. فمثلاً، إذا كان أعضاء جماعة يرون أن أحد أعضاء جماعتهم، وليكن اسمع علياً مثلاً، رجل يمكن الاعتماد عليه والثقة به، فإن هذا الرأي يوجه ويحدد نوع تعاملهم مع علي، وغالباً ما يحدد هذا العامل نوع السلوك من جهة علي كما يحدد نوع التفاعل من جهة الجماعة. فالجماعة تعامله على أنه ثقة، وهو يسلك تجاه الجماعة وفق توقعاتها منه ولا شك أن هذا يعطي تقديرات الأقران صفة أقرب إلى الواقعية.

من الممكن أن تعطينا الملاحظات المنظمة وصفاً أكثر دقة للسلوك أو الأداء المميز إذا فورنت بأساليب التقرير الذاتي أو الأحكام والتقديرات التي يصدرها الآخرون عن الشخص، ذلك لأنه في التقرير الذاتي من المحتمل أن يتدخل للتذكر الانتقائي، بل أنه يحدث أيضاً عندما تستعمل أساليب التقدير

والحكم المختلفة، وكذلك نجد أن الملاحظة في هذه الأساليب الأخيرة ليست قائمة على الاختيار المنظم للمواقف التي يلاحظ فيها السلوك، بل إنها تقوم على ملاحظة المواقف كيفما تقع وقوعها، وفي هذه الأساليب أيضاً لا يلاحظ الشخص في كل المواقف. لهذا قلنا إن للملاحظة المنظمة تعطي وصفاً أدق للسلوك المميز.

ومن السهل أن نلاحظ سلوك مفحوص ما في ظروفه العادية، وهذا ما يسمى بالملاحظة الميدانية، هذا النوع من الملاحظة يكون أكثر ملائمة في بعض الدراسات من الملاحظات المقلنة، فكثير من الباحثين يرى أنه من المستحيل أن نفهم الشخصية إلا إذا لاحظنا المفحوص وهو يستجيب إلى الظروف التي لها مغزى بالنسبة إليه، لأن المثيرات المختلفة لها معان ومغاز مختلفة بالنسبة إلى الأفراد المختلفين. وعلى هذا فهم يرون أن المواقف المقلنة، حيث تلاحظ ردود أفعال الأشخاص لموقف بعينه، تبرز أنماط السلوك الهامة كما تبرزها الظروف العادية، غير المتشابهة، التي يعيش فيها الفرد. إلا أن المشكلة هنا تكمن في كيف نرى قديراً كافيّاً من سلوك الفرد، والحصول على تسجيلات لملاحظات السلوك يمكن الاعتماد عليها. أي أن المشكلة هنا ذات شقين: الأول، عينة السلوك وما يرتبط بها من صعوبات، والثاني، عيوب أو أخطاء من يقوم بالتسجيل، ولنتكلم عن كل منهما بشيء من التفصيل.

لختيار عينة السلوك

لكي نعرف السلوك أو الأداء المميز لشخص ما لا بد أن نعرف ما يميز تصرفه من موقف بعينه، ولكن المواقف تتغير من يوم إلى يوم ومن لحظة إلى أخرى، والطريقة الوحيدة لنطمئن إلى أن ما نلاحظه هو السلوك المميز هي أن

ندرس هذا الشخص في مواقف ومناسبات عديدة، وهو أمر مجهد ومكلف. فالواقع أنه لا بد أن نصل إلى نوع من التوفيق بين الاختيار الدقيق لعينة السلوك والاقتصاد في الإنفاق والجهد.

ولنه من العسير جداً أن نقوم باستنتاجات عن فروق بين الأفراد من ملاحظة سلوكهم في مواقف معينة، لأنه لا يمكن ملاحظة فردين في موقف واحد بعينه، وحتى إذا كان للموقف واحد في ظاهره فإن الظروف السابقة الخاصة بكل منهما تجعلهما يسلكان سلوكاً مختلفاً، ولهذا فإن أقصى ما يمكن أن تقدمه لنا الملاحظات المقارنة هو بيان كيف يسلك أشخاص مختلفون في ظروفهم الحالية، ولكن هذه الملاحظات لا تضمن لنا أن تستمر هذه الفروق عندما تتغير طبيعة هذه الظروف.

ومن أفضل الأساليب التي تؤدي إلى المقارنة الدقيقة ما يسمى بالعينات الموقوتة وفي هذا الأسلوب يخطط مسبقاً جدول أو قائمة للملاحظات ويصمم هذا الجدول بشكل عشوائي يضمن أن يلاحظ كل مفحوص في ظروف متشابهة مع الآخرين بحيث تمكن المقارنة. فمثلاً عند دراسة الاتصالات الاجتماعية للأطفال ما قبل المدرسة وضع جدول لملاحظات لمدة دقيقة واحدة، وبعد ملاحظة الطفل لمدة دقيقة كتب الملاحظة تقريراً كاملاً عما لاحظته من تفاعلات الطفل الاجتماعية. ولوحظ الأطفال وفق نظام مرسوم من قبل وكان يعدل من يوم إلى يوم، وفي خلال الدراسة كان كل طفل يلاحظ عدد مرات متساوية في خلال الخمس دقائق الأولى من ساعات اللعب الحر، ثم الخمس دقائق للثانية وهكذا... وللعينات الموقوتة قصيرة الأجل والموزعة توزيعاً حصناً عدة مزايا، فالصورة المتراكمة من الملاحظات فيها تكون أكثر تمثيلاً للسلوك من نفس القدر من

المعلومات التي يحصل عليها من ملاحظات من مرات أقل وذات فترات طويلة، كما أن أخطاء الذاكرة تقل هنا لأن الملاحظ يقوم بالتسجيل أثناء وبعد الملاحظة مباشرة. والعينات الموقوتة أكثر مناسبة لتسجيل حقائق معينة يمكن أن تصاغ في صورة رقمية، مثل عدد مرات الاتصال الاجتماعي بالآخرين.

ومن الممكن الحصول على عينة أكثر اتساعاً بأسلوب (السجل اليومي) الذي أنشأه باركر وزملاءه. كان هدفهم هو دراسة نمط حياة الطفل ومع الاهتمام الخاص بالمجالات التي يتحرك فيها.

والسجل اليومي مزايًا أن يعطي صور عن تتابع النشاط، ولكن له عيوبه، إذا كان الملاحظ معروفاً ومكشوقاً تأثر السلوك الذي يلاحظه، والاعتماد على يوم كامل بعينه لا يكفي للحصول على معلومات مميزة كاملة عن المنحوس.

ومن المهم أن نلاحظ بالنسبة إلى عينة السلوك الملاحظة عن طريق العينة الموقوتة ما يلي:

1- أن الاستجابات التي بعدها الملاحظ متشابهة أو واحدة قد يكون لها في الحقيقة معان مختلفة. وأن المواقف التي تبدو له متشابهة قد تستدعي استجابات متباينة.

2- أن الاستنتاجات المشتقة من موقف واحد، حتى ولو كانت على أساس تراكم ملاحظات، تكون، أي الاستنتاجات صحيحة بالنسبة إلى هذا الموقف فقط.

3- أن الملاحظات المطلوبة والمناسبة للحصول على معلومات ثابتة ويمكن الثقة بها، يعتمد هذا العدد على المشكلة، إلا أنه بصفة عامة نجد أن عدداً كبيراً من الملاحظات القصيرة أفضل من عدد قليل من الملاحظات التي تعتمد على فترات طويلة.

عيوب من يقوم بالملاحظة

الواقع أن بعض عيوب الملاحظين قد يرجع إلى طبيعة المواقف الاجتماعية ذاتها التي تكون موضوع ملاحظة حيث تحول درجة تعقدها الكبيرة دون التقرير الشامل للمستغرق لها. أما العيوب الأخرى فنرجع إلى الملاحظة نفسه.

وإذا كانت أخطاء الملاحظ مجرد حذف أو إغفال عشوائي لبعض ما يلاحظ فليس لهذا وزن كبير، أما ما يجب الاهتمام به فهو الأخطاء المنظمة، كان يبلغ باستمرار وبصفة مطرودة في تأكيد بعض أنماط الأحداث، كما يغفل الإشارة إلى بعضها الآخر.

عندما يلاحظ مراقبون موقفاً واحداً فإنهم يعطون تقديرات بدرجة كبيرة، ذلك لأن كل ملاحظ عنده حساسية لأنماط معينة من السلوك أكثر من غيرها والوضع الأمثل هو أن يبني الملاحظ انطباعه على كل تصرف يمكن أن يكشف أو يعكس خصائص أساسية، ولكن عندما يركز من يقوم بالملاحظة والمراقبة على شيء بعينه فإنه بالضرورة لا بد أن يغفل ملاحظة أشياء أخرى.

رئمة عيب آخر وهو أن الملاحظين يفسرون ما يرون، فإذا قام الملاحظون بتسجيل الحقائق الموضوعية، فقد يصل آخرون إلى تفسيرات مختلفة

تماماً، ولكن الناس يميلون إلى تفسير ما يرون، وعندما يفسرون فإنهم يميلون إلى تجاهل الحقائق التي لا تتفق مع تفسيرهم بل أنهم قد يخترعون حقائق لتكمل تفسيرهم.

ومما يقل من آثار هذه العيوب أن يتم التسجيل بشكل منظم، وأن يكون كلما أمكن، في صورة وحدات سلوكية يمكن عدّها فمتلاً، من الممكن أن نصف مدى انتظام عمل مصنع في عملهم عن طريق التسجيل لزماني للحظات التي يكونون فيها في عمل فعلي، وللحظات ينشغلون فيها عن العمل للحصول على الأدوات، أو زيارات زملائهم، أو يتعرضون فيها لمؤثرات التشويش. وحتى إذا كان السلوك للملاحظ على درجة كبيرة من التنوع بحيث يصعب وضعه في صورة وحدات سلوكية فمن الممكن أن تحدد أقسام السلوك وتعد قبل ويكون عمل الملاحظ هو أن يصنف السلوك الذي يلاحظه وفق هذا التقسيم، ويضع علامة في القسم المناسب لكل أسلوب يلاحظه كما هو الحال في أقسام بيلز.

وعلى الرغم من فائدة التسجيل والحد الموضوعي لوحدات السلوك وملائمته للبحوث العملية فإن فائدته محدودة في التوجيه والإرشاد النفسي.

أما السجلات القصصية أو الوقائعية فإنها، وإن كانت تنقصها دقة الموضوعية، فإنها تعطي وصفاً حياً للشخص، لأن الملاحظ فيها يوجه اهتمامه إلى أي سلوك يبدو جدير بالملاحظة بدل أن يركز على سمات معينة للمفحوصين كعلم إلا أن الملاحظ في السجل القصصي يصف بدقة ما يلاحظه فاصلاً بين الوقائع وتفسيراتها. وعندما تتراكم هذه السجلات القصصية فإنها تعطي صورة أكثر غنى من أي وسيلة أخرى على هذه الدرجة من البساطة. ومسؤولية الملاحظ الكبرى هنا تتركز في اختيار الوقائع التي تستحق التقرير، وفي أن يتحرى الموضوعية بقدر الإمكان.

قياس الشخصية وسائل التقرير الذاتي

وملأ التقرير الذاتي

أشرنا فيما سبق إلى تصميم المقاييس النفسية بصفة عامة إلى قسمين رئيسيين هما: مقاييس الأداء للمميز، ومقاييس أقصى الأداء، وعرفنا أن في النوع الأول من هذه المقاييس يكون الهدف من القياس هو معرفة ما يقوم به الشخص عادة، بينما النوع الآخر من المقاييس يكون الهدف هو معرفة أقصى ما يمكن أن يقوم به للشخص في ظروف استثارة ودافعية غير عادية.

ويمكن أن نضع تحت القسم الأول، أي مقاييس الأداء للمميز، اختبارات الشخصية، واختبارات الميول والاتجاهات بينما تدخل اختبارات القدرة بأنواعها المختلفة تحت القسم الثاني وهو اختبارات أقصى الأداء. وفي هذا الفصل نتكلم عن قياس الشخصية باعتباره قياساً للسلوك أو الأداء للمميز. وسوف نتبع، بصفة عامة، للتقسيم الذي حدده كرونباك لأنواع مقاييس الشخصية، فندرجها تحت هذه الأقسام الثلاثة:

- 1- قياس الشخصية عن طريق التقرير الذاتي، وهي ما سيتناوله الفصل الحالي.
- 2- قياس الشخصية عن طريق أحكام الآخرين. والملاحظات المنظمة، وهي ما سيتناوله الفصل التاسع.
- 3- قياس الشخصية عن طريق اختبارات الأداء وهي ما سيتناوله الفصل العاشر.

قياس الشخصية عن طريق التقرير الذاتي:

على الرغم من أن تعبير السلوك المميز الذي تكررت الإشارة إليه يحمل طابعاً (سلوكياً) فإن المذهب السلوكي لم يكن هو الذي أسهم في نشأة وتطوير مقاييس للشخصية وإنما يرجع الفضل في هذا إلى الاتجاه الظاهراتي (الفيثونولوجي)، فهذا الاتجاه يتميز بأنه يركز على العالم كما يدركه للشخص، أو على العالم الذاتي للشخص، بما في ذلك إدراك الفرد لذاته، وعلى أن السلوك إنما يتحدد، ولا يمكن أن يفهم إلا في ضوء هذا العالم الذاتي الداخلي، كما يذهب هذا الاتجاه إلى أن غالب مشكلات التوافق التي يتعرض لها الفرد ترجع إلى إدراك الشخص للأحداث والأشخاص وليس لواقع هذه الأحداث أو الأشخاص في ذاتهم. وبالتالي فإن العلاج النفسي، وفق هذا الاتجاه يجب أن يركز على تصحيح إدراك الشخص.

نشير إلى المدرسة الفيثونولوجية في بداية كلامنا عن قياس الشخصية لأن أول مقاييس أو استفتاءات للشخصية كانت قد وضعت بهدف قياس هذا العالم الداخلي الذاتي من الإدراك والشعور. ففي العقد التاسع من القرن التاسع عشر وضع جالتون هذه المقاييس عندما احتاج إلى طريقة يمكن أن يطبقها على عدد كبير من مفحوصيه أثناء دراساته للصور العقلية، ثم استعمل هذا النوع من الاستفتاءات بعد ذلك ستالي هول في دراساته الموسعة عن نمو المرافقين ولكن الاستفتاء كانت له وظيفة خالصة بالنسبة إلى كل من جالتون وستالي هول، فبالنسبة إلى جالتون كان التقرير الذاتي، يستعمل على أنه الوسيلة الوحيدة الممكنة للحصول على معلومات عن أمور وأحداث في العقل للمفحوص، في داخل رأسه، بينما كان التقرير الذاتي بالنسبة إلى ستالي هول يستعمل بهدف

تجنب الجهد الشاق والوقت الطويل الذي يمكن أن يستهلك في الملاحظة المباشرة لسلوك المراهقين.

لما أول استفتاء وضع بقصد إيلس الشخص فقد كانت قائمة وود ورث للبيانات الشخصية. وظهرت خلال الحرب العالمية الأولى عندما ظهرت الحاجة إلى تمييز أولئك الجنود الذين عندهم قابلية للانهايار أثناء القتال، ولم يكن يتوافر عدد كف من المتخصصين لتحقيق هذا الهدف. لهذا وضع وودورث قائمة من الأعراض التي يهتم الفاحص بالتعرف عليها لاستبعاد من تظهر عندهم، وعندما أعطيت القائمة ميزت بالفعل أولئك الذين قرروا وجود عدد كبير من الأعراض عندهم ثم خضعوا لاختبار أكثر دقة، وكانت لاختبار البيانات للشخصية هذا قيمة كبيرة في ذلك المجال.

ويعتبر اختبار وودورث لجد الأكبر لعدد كبير من استفتاءات للتوافق، وهي التي تتكون أساساً من قوائم من المشكلات أو الأعراض أو الشكاوى التي على المفحوص أن يختارها. إلا أن مثل هذه القوائم لا تدعى أنها تعطي وصفاً دقيقاً عن الشخصية وهي غالباً ما تعطي درجة واحدة تمثل مستوى للتوافق. وهي تتكون من عناصر يمكن أن يميز بين أشخاص معروف أنهم غير متوافقين وبين آخرين معروف أنهم عاديون. والاستعمال الرئيس لمثل هذه المقاييس هو التعرف على أولئك الأشخاص الذين يحتاجون إلى رعاية وإرشاد نفسي، وعلى هذا يمكن أن ننظر إليها على أنها أدوات فرز تساعد في تحديد الأشخاص الذين يختارون أعراضاً وصفات وانتقادات تتصل بهم شخصياً.

ولا يمكن اعتبار استفتاءات التوافق مقاييس دقيقة لسمات محددة، فالمعلومات التي تمدنا بها معلومات ليس لها صفة العمق. لهذا فإنه في الفترة

بين 1920-1945، تغيرت النظرة إلى مثل هذه الاستفتاءات وذلك بتأثير المدرسة السلوكية واستكراها للتأمل للباطني الذاتي الذي هو أساس اختبارات للتوافق، فيجب أن يكون المقياس بديلاً لملاحظة السلوك، وعلى هذا يجب أن تؤكد أسئلته بشكل أكبر على ما يفعله الشخص لا على ما يشعر به أو يفكر فيه. وعلى هذا اتسعت مجالات الاختبارات بحيث تصف أكبر عدد ممكن من جوانب الشخصية، وتلخص الاستجابات بحيث تصف أكبر عدد ممكن من جوانب الشخصية، وتلخص الاستجابات بإعطاء درجات لعدد من السمات أو الأنماط الاستجابات، ذلك لأنه كان ينظر إلى الشخصية في ذلك الوقت على أنها مجموعة أو حزمة من العادات، وكان يحكم على الشخص حسب قوة سمات مثل: الثقة بالنفس، الإصرار، سهولة تكوين صداقات... إلخ والسمة القوية هي السمة التي تصف استجابة يقوم بها الشخص عادة وبكثرة.

وقد كانت قائمة السمات أو الخصام السلوك التي تعطى لها درجات في الاختبارات تختار بشكل قطعي ودون خطة معينة في الاختبارات الأولى، فمثلاً سمة مثل الثقة بالنفس كان مصدرها الخبرة اليومية، بينما تشتق سمة أخرى مثل الانطواء من نظرية من نظريات الشخصية. وقد ظهرت عشرات من الاختبارات بهذا الشكل، تأخذ اللاحقة منها من السابقة عليها بضعة عبارات أو أسئلة وتضيف إليها وتصممها بحيث توضع في تجمعات سمات مختلفة. إلا أن أكثرها شيوعاً هو اختبار بيرنرويتز للشخصية.

وتعطى فيه للدرجات للدلالة على النزعة العصابية، الاكتفاء للذاتي، الانطواء، والسيطرة. ويتكون الاختبار من (25) سؤالاً يجيب عليها المفحوص بنعم أو لا. وقد وضعه بيرنرويتز على أساس أن سلوك الفرد في مواقف واحد

قد يكشف عن سمات شخصية مختلفة، وعلى هذا يمكن أن ينظر إلى كل سؤال من أسئلة الاختبار على أنه يقيس أكثر من سمة من السمات، وبالتالي يمكن أن تستخدم الاختبار الواحد للكشف عن أكثر من سمة تبعاً لطريقة معالجة الاستجابات التي تحصل عليها.

وقد كانت دراسة فلانجان لاختبار بيرنرويتز بداية لاستعمال درجات السمات كما تحدها القواعد الإحصائية. لقد اعتقد فلانجان أن السمات لا يمكن أن تطلق عليها أسماء خاصة بها إلا إذا كانت الارتباطات فيما بينها منخفضة. وعلى هذا قام بحساب معاملات الارتباط لدرجات اختبار بيرنرويتز من تطبيقه على 305 مراهق، ووجد أن هذه السمات ليست مستقلة. فقد وجد أن الخطوات لا يختلف إلا قليلاً عن اللزعة العصبية، وعندها طبق التحليل العاملي وجد عاملين هما: الثقة بالنفس والميل الاجتماعي، يمكن أن يفسر البيانات التي نحصل عليها من قياس السمات الأربعة الأصلية، ولشأن مفاتيح خاصة بقياس العاملين أو المسمتين اللتين وصل إليهما أي ثقة بالنفس والميل الاجتماعي. وقد ارتبطت درجات هاتين المسمتين ارتباطاً ضعيفاً جداً، وعلى هذا أصبح من الممكن أن نقول أنهما يمثلان جانبين مستقلين من التقرير الذاتي.

بعد أن حدث تطور آخر له مغزاه، ذلك هو إخضاع نظريات الشخصية تماماً لبحث إحصائي عن أبعاد يمكن أن تلخص الشخصية. فقد أدت الارتباطات الداخلية بين عناصر الاختبارات إلى أن اقترح جيلفورد مثلاً، أن الخطوات يمكن أن تنقسم إلى خطوات اجتماعي (ج)، وخطوات فكري (ف)، ولكتلب (ك)، والتقلب المزاجي (ت)، وضبط للنفس (ض). وعلى هذا أنشأ الاختبار الذي أسماه: اختبار العوامل (ج - ف - ك - ت - ض).

ثم أضاف بعد هذا ثمانية جواذب أخرى للشخصية. ولم تكن مقاييس جيلفورد مستقلة، ولهذا أعاد بعض الباحثين تنظيم هذه المقاييس في أنماط تجعلها أكثر كفاءة. من هؤلاء مثلاً ترستون الذي استطاع أن يفسر كثيراً من المعلومات في مقاييس جيلفورد الثلاثة عشرة بسبعة عوامل فقط وأعاد تسميتها بحيث أصبحت هكذا للتأمل، الميل الاجتماعي، الاتزان الانفعالي، الحيوية أو الذكورة، السيطرة، للنشاط الاندفاع.

وهكذا نجد أن واحداً من علماء النفس يصنف عناصر الاختبار التي تقيس السمات بشكل دقيق تفصيلي، بينما يأتي التالي ليضعها في مجموعات صغيرة، ثم يأتي آخر فمعيد تقسيم المجموعات بطريقة جديدة، ويعطي كلاً منها أسماء جديدة للعوامل التي يصل إليها، وبهذا لا تنتهي هذه اللعبة، كما يطلق كروبناك، وفي رأيه أنه ما لم ترتبط قوائم السمات بنظرية محددة وبمحات خارجية، فإن عملية الاختيار لا يمكن إلا أن تكون عملية تفضيلية شخصية. وليس هناك اتفاق في الوقت الحاضر بين أصحاب التحليل العاملي، بالنسبة إلى عدد للعوامل التي تم تحديدها بشكل دقيق، أو لفضل تنظيم لها، أو لنسب الأسماء لها.

أما الاتجاه الأحدث في قياس الشخصية بصفة عامة فهو تحديد المفاهيم والتكوينات النظرية على أساس نظرية الشخصية ثم إعداد للعناصر في الاختبارات بصورة تؤدي إلى الحصول على معلومات عن هذه المفاهيم والتكوينات. وقد كان لنظرية يونج عن الشخصية تأثير في الدراسات الأولى عن الانطواء، ولكن هذه النظرية لم تتعد اقتراح بعض عناصر اختبارات وضعت لمجرد المحاولة. أما الآن فهناك دراسات واسعة عن اختبار مايرز - برجز

والتي اشتملت على صوره ومفاتيح تصحيحه صراحة من نظرية يونج. وتوجد مقاييس أخرى تمثل هذا الاتجاه منها مقياس التفضيل الشخصي لإنواردز.

والذي يشتق من نظرية موري عن الحاجات الخمس عشرة، ومن أمثلة هذه الحاجات: الإنجاز (التحصيل)، الخضوع، والنظام، والعرض، والاستقلال الذاتي، السيطرة، التأمل النفسي (والمقصود بهذه الحاجة تحليل دوافع ومشاعر الفرد الذاتية وعند الآخرين)، ويتكون المقياس من 210 زوجاً من العبارات، وعلى المفحوص أن يختار عبارة من كل زوج على أنها أكثر تمثيلاً لخصائص شخصيته، ومن أمثلة هذه الأزواج:

- أ- أحب أن أتحدث عن نفسي مع الآخرين.
- ب- أحب أن أعمل لهدف معين حددته بنفسى.
- ج- أشعر بالاكتراب عندما أفضّل فى أمر ما.
- د- أشعر بالاضطراب عندما أتحدث إلى جماعة.

ويهدف هذا المقياس إلى تقدير القوة النسبية الخمسة عشر متغيراً المرتبطة بالحاجات.

ويمكن تقويم الخمسة عشر درجة المرتبطة بالحاجات على أساس معيار منوية وثانية لطلاب وطالبات الجامعة، وهذه للمعايير تقويم على عدد 749 طالباً، و 760 طالبة موزعين على جميع أنحاء الولايات المتحدة، وقد أضيفت معايير منوية للكبار ينتمون إلى بيئات حضرية وريفية.

أما عن ثبات المقياس فقد ذكر دليل الاختبار معاملات ثبات لمقاييسه الفرعية للخمس عشرة معتمدة على إعادة الاختبار بعد أسبوع واحد تتراوح بين

0.74، 0.88، بينما معاملات الثبات المعتمدة على التجزئة للنصفية تتراوح بين 0.60، 0.87.

أما بالنسبة إلى صدق مقياس إدواردز، فعلى الرغم من قلة البيانات الواردة عنها في دليل الاختبار ذاته، فإن أبحاثاً كثيرة مستقلة حول صدقه قد نشرت، إلا أنها أعطت نتائج متعارضة، وتذهب أتمتازي إلى أن مقياس التفضيل الشخصي له مزايا كثيرة إلا أنه محتاج إلى أمرين:

أولاً، مراجعة للتخلص من بعض جوانب الضعف الفنية فيه، وبصفة خاصة فيما يتصل بشكل العناصر وتفسير الدرجة، وبالتالي، دراسات صدق تجرى بدقة مستفيدة من أساليب تطيل أنماط الدرجات مناسبة للدرجات هذا المقياس التي لا تتحدد فيها قوة الحاجة بصورة مطلقة ولكن بالنسبة إلى قوة الحاجات الأخرى عند الفرد، وعلى هذا فإن إطار المرجع في هذا النوع من الدرجات هو الفرد ذاته وليس عينة للتقنين. وقد يكون شخصان متفقان في درجاتهم على المقياس إلا أنهما يختلفان اختلافاً جوهرياً في القوة المطلقة لحاجاتهم.

وحسب ثبات الاختبار بطريقة التجزئة للنصفية، وبتطبيقه على طلاب الجامعة، وكان يتراوح بين 0.34، 0.77 وقد استعمل هذا المقياس في بعض الدراسات المقارنة بين الثقافات وغيرها.

وقد ارتبط في تصميمه بالبحث في نظرية هل سبيل السلوكية، وسوف نعود إليه بعد دراستنا لاختبار مينسوتا المتعدد الأوجه الذي توليه اهتماماً خاصاً.

اختبار مينسوتا المتعدد الأوجه للشخصية

يعتبر هذا الاختبار من أهم الأدوات التي وضعت لقياس الشخصية عن طريق التقرير الذاتي، لهذا سوف نتناوله بشيء من التفصيل. نجد لهذا الاختبار مكاناً بين اختبارات الشخصية تشبه مكان اختبار سترونج بين اختبارات الميول، ذلك لأنه يشترك اختبار سترونج في الأساس التجريبي الذي صمم وفقاً له، وكذلك يشبه من حيث تعرضه لبحث عميق وطويل من جانب واضعيه. وقد ظهر هذا الاختبار في وقت كانت الحاجة إليه شديدة، وكان الاعتماد عليه كبيراً أثناء وبعد الحرب العالمية الثانية التي شهدت توسعاً كبيراً في علم النفس المرضي والعلاج النفسي. أي أن نشأته كانت أصلاً نشأة إكلينيكية وهذا مما ساعد على انتشاره، وكذلك ساعد على انتشاره أيضاً بعض الأمور الفنية المبتكرة فيه ومنذ نشره في عام 1940 ونشر دليله في عام 1943 فإنه استلزم أيضاً من البحوث التي استخدمته.

وكان تصميم الاختبار في الأصل على يد عالم نفسي هو ستارك هاتلوي وطبيب نفسي هو ج. س. ما كيلي وكان الهدف من هذا التصميم هو المساعدة في تشخيص المرضى النفسيين، وتطويل السمات التي تميز للشذوذ النفسي الذي يؤدي إلى المعجز.

يتكون الاختبار من 550 عبارة يسأل المفحوص أن يصنفها وفق أحد أقسام ثلاثة، أي يجيب عنها إما بنعم أو لا، ولا يستطيع أن أجزم أو لا أدري. وفي الصورة الفردية للاختبار كانت العبارات تطبع على بطاقات منفصلة وعلى المفحوص أن يصنفها في ثلاث مجموعات، وبعد ذلك طبعت العبارات في كتيب وتسجل الاستجابات في ورقة الإجابة. وقد صممت للصوتان الفردية والجماعية

للكبار، أي من 16 سنة وما بعدها، إلا أنهما استخدمتا بنجاح مع مراقبين أصغر من هذا السن. وبصفة عامة يفضل استخدام الصورة الفردية وخاصة عدد الأشخاص المضطربين أو من ذوي المستوى التعليمي أو العقلي المنخفض.

وتتنوع عناصر الاختبار المتعدد الأوجه للشخصية وتتسع من حيث محتواها فتشمل مجالات مثل: الصحة، والأعراض الجسمية- النفسية، والاضطرابات العصبية، والاتجاهات نحو الجنس، والسياسة، والدين، والمساكن التربوية والمهنية والعائلية، كما تشمل كثيراً من المظاهر والأعراض العصبية والذهالية المعروفة مثل المخاوف المرضية، والاضطهاد، ونزعات تعذيب الذات وتعذيب الآخرين وغيرها.

وعندما نشر الاختبار في صورته الأولى كان يعطى درجات على تسعة (مقاييس إكلينيكية)، ويتكون كل مقياس منها من عناصر ميزت بين جماعة مرضية محدودة وبين جماعة ضابطة عادية، وعلى هذا فإن هذه المقاييس قد أنشئت بطريقة تجريبية، وذلك عن طريق تقسيم العناصر وفق محك معين وهو التشخيص النفسي المعروف ووفق هذه الطريقة أعدت المقاييس التالية:

- 1- توهم المرض.
- 2- الاكتئاب.
- 3- للهستيريا.
- 4- الانحراف السيكوباتي.
- 5- الذكورة، الأنوثة.
- 6- البارافانيا.
- 7- السيكاينزيا.

8- لفصلام.

9- الهوس الخفيف.

ولما يلي تعريف مختصر بما يدل عليه كل من هذه المقاييس:

- 1- توهم المرض: وهو مقياس للاهتمام الزائد بالوظائف الجسمية، والقلق الذي لا مبرر على الصحة.
- 2- الاكتئاب: ويرتبط هذا المقياس بالشعور بالانقباض والتشاوم والإحساس بحجم الأهمية الذاتية.
- 3- الهستيريا: ويقاس هذا المقياس درجة تشابه المفحوص بالمرضى الذين تظهر عندهم أعراض تتخذ صورة شكوي عامة أو محددة مثل العمى أو اللثال أو القيء أو الاضطرابات المعوية أو الأعراض القلبية.
- 4- الانحراف السيكوبتي: وهو مقياس لدرجة تشابه المفحوص بجماعة السيكلبيين الذين تتمثل المعوية الرئيسية عندهم في نقص الاستجابة الانفعالية العميقة وقلة الاهتمام بمشاعر الآخرين، وعدم المبالاة بالمعايير الاجتماعية.
- 5- الذكورة - الأنوثة: وهو مقياس للذكورة والأنوثة في أنماط الاهتمامات، وقد استخرجت عبارات هذا المقياس نتيجة مقارنة استجابات الذكور من ذوي الاهتمامات الذكرية باستجابات الذكور من ذوي الاهتمامات الأنثوية، أما الفروق بين استجابات الذكور والإناث فقد درسها واضعاً الاختبار إلا أن هذه الفروق لعبت دوراً صغيراً نسبياً في هذا المقياس. وتدل الدرجات المرتفعة على هذا المقياس على غلبة أنماط الاهتمامات الخاصة بالجنس الآخر.

6- البارثوليا: وهذا مقياس للتشكك للزائد، والحساسية المفرطة، وهو اجس الاضطهاد.

7- السيكتينيا: وهو مقياس للمخاوف المرضية والسلوك القهري، الذي قد يكون صريحاً مثل التكرار للزائد لغسل اليدين، أو ضملياً كما يتمثل في عدم القدرة على التحرر أو الهروب من فكرة متسلطة.

8- الفصام: ويكشف هذا المقياس عن التفكير أو السلوك الخاطي للشاذ والذي لا يتفق مع الاتصال بالواقع وفهمه.

9- الهوس الخفيف: وهو مقياس للنشاط للزائد في الفكر والعمل، وعدم التركيز الطويل، والانشغال بحد كبير من الأمور والمشروعات في وقت واحد.

وتعطى بيانات جماعة التقنين من العالدين درجة معيارية (ت) يمكن تحويل الدرجات الخام إليها بحيث تضبط لنتائج على بطاقة صحيفة نفسية وتتركز الأهمية الأولى على الدرجات التي تزيد عن 70 (والدرجة 50 هي متوسط جماعة المرجع)، وهذه الدرجة الفاصلة، أي 70، هي درجة مختارة على أساس قطعي اعتصالي، أي ليس لها أسس أو تبرير نظري، وعلى أي حال نجد أ، مفسري نتائج الاختبار يميلون إلى الاهتمام بكل الارتفاعات، أو لقسم، في بطاقة للصحيفة النفسية، سواء كانت هذه الارتفاعات تتخطى خط (70) درجة أو لا تتخطاه.

مقاييس الضبط

ومن العلامات المميزة والخاصة بمقاييس ميلسوتا المتعدد الأوجه استعماله لأربعة مقاييس تسمى مقاييس للصدق، وللصدق هنا معنى الضبط لأنها

في الواقع تمثل ضوابط للإهمال، وسوء الفهم، والتحليل، ولتأثير أنماط الاستجابات والاتجاهات نحو أخذ الاختبارات، وهذه المقاييس هي:

1- الأول ؟: وتمثل درجة هذا المقياس مجموع العناصر التي أجاب عنها المفحوص بالآلا أستطيع أن أجزم، ومن المرغوب فيه أن تكون هذه الدرجة أقل ما يمكن.

2- الثاني L: ويمثل للكنب، ويعتمد على مجموعة من العناصر التي تجعل للمفحوص يبدو في شكل مقبول اجتماعياً، ولكن ليس من المحتمل أن تكون الإجابة عنها صادقة في الاتجاه المرغوب فيه، ومن أمثلة هذا العنصر (لا أقول للصدق دائماً).

وعلى الرغم من أن الإجابة للصحيحة المعتادة تكون نعم، إلا أن الإجابة المقبولة اجتماعياً تكون لا، ومن المفترض أن للشخص الذي يريد أن يظهر نفسه في صورة مقبولة يحصل على درجة مرتفعة في المقياس L عن طريق تحريف استجاباته لعبارات المقياس، وتتل للدرجات العالية على هذا المقياس على أن الإجابات لا يمكن الثقة بها، ولكنها لا تكل على أن المفحوص يعتمد الكذب.

3- الثالث F: وهو مقياس للخطأ، ويتكون من العناصر التي ندر أن أجيب من مجموعة للتقنين في نفس الاتجاه الذي لجابه فيها للمفحوص، وعلى الرغم من أن هذه العناصر لا يمكن أن تمثل نمطاً معيناً من التشنوذ النفسي فإنها تعبر عن سلوك غير مرغوب فيه، وعلى هذا فإنه من غير المحتمل في الواقع أن يعبر مفحوص واحد عن كل أو معظم هذه الأعراض. وترتفع للدرجة على هذا المقياس عندما لا يستطيع

المفحوص أن يعطي إجابة مميزة لمسبب من الأسباب كلن لا يستطيع القراءة أو الفهم بدرجة مناسبة، أو الإهمال، سواء عن قصد أم غير قصد، كما قد يرجع ارتفاع الدرجة إلى الخطأ في التصحيح. وبصفة عامة كلما زادت الدرجة على هذا المقياس زاد احتمال أن بعض العوامل تدخلت لتقلل من صدق الصحيفة للنفسية (أو البروفيل).

4- الرابع K: ويمثل مقياس التصحيح، وفي هذا المقياس استعمال للتركيبة أخرى من العناصر المختارة، وتعطى درجة هذا المقياس دلالة على الاتجاه نحو أخذ الاختبار، وهي متصلة بكلا المقياسين السابقين I، F، ولكن يعتقد على أنها أكثر خفاء منهما وتدل الدرجة العالية على هذا المقياس على اتجاه دفاعي أو على محاولة (المخادعة في الاتجاه الحسن) أو التحريف المقصود نحو الطرف السوي، كما تدل للدرجة المنخفضة على الصراحة المفرطة، وقد لذات، أو المحاولة المتعمدة للمخادعة في الاتجاه السيئ أو الميل إلى إظهار النفس بالمظهر غير السوي.

ومن المهم أن نشير إلى أن استعمال مقياس الضبط الأربعة هذه ليس مقناً تماماً ولكنه يترك جزئياً لحكم مستعمل الاختبار، وكذلك تشير إلى أن هذه المقاييس تخضع باستمرار لإعادة التحديد والمراجعة في ضوء للبحوث الجديدة.

كانت تلك هي المقاييس الإكلينيكية التسعة الأساسية ثم المقاييس الأربعة الأخرى الخاصة بضبط الاختبار، إلا أنه منذ نشر الاختبار أنشئ عدد كبير جداً من المقاييس الجديدة بواسطة باحثين آخرين لم يشاركون في وضع المقاييس الأصلية، من هذه المقاييس الجديدة مقياس عن الانطواء الاجتماعي، وقد أضيف

إلى المقاييس التسعة الأصلية وهو مضمن في الاختبار ورقمه (صفر). ويقس هذا المقياس النزعة إلى الانزواء والابتعاد عن الاتصال الاجتماعي بالآخرين.

وبهذا فإن الاختبار في صورته الحالية يعطي (14) درجة على الأربعة عشر مقياس، منها التسعة الأصلية، ومقياس الانطواء الاجتماعي، ثم مقياس لضبط الأربعة.

ومعايير المجموعة الضابطة الأصلية والتي تتكون من حوالي (700) فرداً مذكورة في صورة درجات تقيية. أو درجات معيارية بمتوسط (50) والانحراف المعياري قدره (10). وتستعمل هذه الدرجات المعيارية في رسم الصحيفة النفسية. وأي درجة تعادل (70) أو تزيد عنها، أي تقع فوق المتوسط بدرجتين لانحراف معياري، ينظر إليها بصفة عامة على أنها تميز انحرافاً مرضياً ومما يجب ملاحظته أن المعنى الإكلينيكي لنفس الدرجة قد يختلف من مقياس إلى مقياس آخر. فمثلاً الدرجة (75) على مقياس توهم المرض وعلى مقياس للفصام لا تكل على نفس الدرجة من حدة الانحراف.

أما تفسير الدرجات المنخفضة على الصحيفة النفسية للاختبار فليس على درجة كافية من الوضوح. إذ يعتقد الإكلينيكيون أن وقوع درجات أقل من (50) كثيراً قد يكون له مغزى تشخيصي، ولكن لم توضع بعد تفسيرات منظمة لهذا.

على أي حال، فإن هناك قديراً من نتائج للبحوث التي أجريت على هذا الاختبار يوضح أنه بصفة عامة كلما زاد عدد وحجم الدرجات المنحرفة على الاختبار، زاد احتمال أن يكون للشخص مصاباً باضطراب شديد.

ومن الواضح أن الاستعمال الرئيسي للاختبار المتعدد الأوجه هو التشخيص المتمايز وعندما يستعمل في هذا المجال لا بد أن يمارس مستعملة منتهى الجنور ولا يلجأ إلى التفسير الحرفي للمقاييس الإكلينيكية فيه، لأن الأمر أكثر تعقيداً من تلك الأسماء المعطاة للمقاييس الأصلية، فليس معنى الدرجة المرتفعة على مقياس للفصام مثلاً أن للشخص مصاب بالفصام، فقد يحصل شخص عادي على نفس الدرجة، بينما قد يحصل المصابون بالفصام فعلاً على درجات مرتفعة على مقاييس أخرى. ولكي يستبعد احتمال سوء التفسير للدرجات على المقياس الواحد فإن أرقماً من صفر إلى 9 حلت محل أسماء المقاييس في طبعات أخيرة من الاختبار. ولكي يسهل النظر في نمط الدرجات أو الصحيفة النفسية ولحصها وضع نظام ترميز عددي للصحيفة النفسية. وفي هذا النظام يدل تتابع وترتيب أرقام المقاييس، بنظرة سريعة، على للنقط المرتفعة والمنخفضة في الصحيفة النفسية لشخص ما. فمثلاً الرمز 2-94 يدل على درجة مرتفعة على المقياس رقم (4) (ب د) وعلى المقياس رقم (9) (م 1) بترتيب متناقص، أي أن الدرجة على المقياس الأول أعلى منها على المقياس الثاني، كما يدل هذا الرمز على درجة منخفضة على المقياس رقم (2) (د).

والواقع أن أسماء المقاييس التي استحدثت كانت تستند إلى تقسيم إكلينيكي يقوم على أساس نظري مشكوك فيه على الرغم من شيوعه وقت وضع الاختبار، فالتقسيمات المستعملة والتي وضعت لها مقاييس الاختبار كانت محل شك ومراجعة على حتم النفس المرضي، أما في المجال العلمي الواقعي فإن هذه التقسيمات لم تعد مقبولة.

ومن دواعي الحذر في تفسير درجات المقاييس أن بعض هذه المقاييس لا تتمتع بثبات مرتفع. وتعتبر هذه نقطة ضعف في الاختبار. لأن فاعلية أي تحليل للصحيفة النفسية تقل بسبب أخطاء الصدفة في الدرجات التي تعتمد عليها هذا التحليل. فإذا كانت درجات المقاييس المفردة غير ثابتة ومتداخلة أي ترتبط فيما بينها ارتباطاً عالياً فإن كثيراً من الاختلافات فيما بين الدرجات ربما تكون ناتجة عن الصدفة.

وتتراوح معاملات ثبات الاختبار للناتجة عن إعادة الاختبار، بالنسبة إلى الراشدين الأسوياء وغير الأسوياء بين 0.50، 0.9 وكانت الفترة بين الاختبار وإعادة نتراوح بين بضعة أيام إلى عام. إلا أي هناك دراسة أجريت على طلاب الجامعة كانت معاملات الثبات فيها عموماً أقل من السابقة، كما أظهرت للدراسة نفسها تنوعاً كبيراً بين مقاييس الاختبار كانت تتراوح بين 0.05 لمقياس (ب) 1، 81 لمقياس (ب ت).

وثمة عيب آخر في الاختبار يتصل بمدى تمثيل عينة التقنين. فالدرجات المعيارية التي اشتقت منها تطبيقات الصحائف النفسية موضوعة في ضوء جماعة ضابطة تتكون من 700 راشد الذين اختبروا في التقنين الأصلي للاختبار. وتبدو هذه العينة قاصرة جداً إذا قورنت بعينات التقنين الأخرى الأكثر شمولاً والتي استعملت، مثلاً، في تقنين اختبارات القدرة وقد أكتت أبحاث كثيرة ضرورة أن تكون المعايير متنوعة تنوعاً كبيراً في جامعات أصل مختلفة، ومن هذه الأبحاث ما وضع أن متوسطات درجات طلبة الكليات كانت فوق (50) بشكل مطرد على بعض المقاييس، كما بينت دراسة أخرى أن 39% من طلبة

الكليات حصلوا على درجات فوق (70) على مقياس أو أكثر من مقاييس الاختبار.

ويتأثر الاختبار المتعدد الأوجه، مثل سائر اختبارات الشخصية، بالاختلافات الثقافية العامة بين المجتمعات بل والمستويات الثقافية الفرعية داخل المجتمع. وقد أظهرت دراسات كثيرة أجريت في بلاد مختلفة ارتفاعاً ملحوظاً في بعض مقاييس الاختبار عندما وضعت الصحائف النفسية على المعايير الأمريكية الأصلية. وإن تفسير مثل هذه الاختلافات في ثقافة المجتمع أو ثقافته الفرعية يتطلب معرفة بخصائص هذه الثقافات وظروفها، فقد يكون الاختلاف في درجات الاختبار المتعدد الأوجه راجعاً إلى اختلاف الأفراد في تفسير عبارات الاختبار أو تعليمات الاختبار، فمثلاً قد يرجع الارتفاع في بعض المقاييس عند بعض الجماعات إلى اعتقادات متوارثة نحو التواضع وتقليل شأن الذات. وكذلك تتأثر الدرجة بالاختلافات الثقافية يتصل بالسلوك الذي يعتبر مقبولاً اجتماعياً. إلا أن الدرجات المرتفعة على مقاييس هذا الاختبار، بالنسبة إلى بعض المجتمعات والجماعات، قد تدل على وجود مشكلات انفعالية حقيقية يمكن إرجاعها إلى أساليب تنشئة الطفل، أو الصراعات في الأدوار الاجتماعية، أو غير ذلك من أساليب الاضطراب الانفعالي.

اختبارات مشتقة من اختبار الشخصية المتعدد الأوجه

إلى جانب حركة الليحوت الضخمة التي حركها الاختبار المتعدد الأوجه فقد كان له في ذاته خصوصية عالية من حيث اشتقاق مقاييس جديدة منه أو تأسيس اختبارات وتطويرها معتمدة عليه.

فن المقاييس الجديدة الكثيرة التي استخرجت منه عن طريق استعمال مفاتيح تصحيح جديدة لنذكر ما يلي:

- 1- مقياس للتعصب ويهدف إلى قياس العوامل السيكولوجية المرتبطة بالتعصب ضد جماعات الأقلية.
- 2- مقياس المكائنة الاجتماعية - الاقتصادية ويقس بعض العوامل السيكولوجية المرتبطة بالمكائنة الاجتماعية - الاقتصادية للشخص.
- 3- مقياس الانحراف الجنسي.
- 4- مقياس قوة الأداة.

ويعتبر مقياس تايلور للقلق الظاهر الذي سبقت الإشارة إليه، من المقاييس التي اعتمدت على اختبار الشخصية المتعدد الأوجه. وقد استعمل هذا المقياس استعمالاً واسعاً في البحث وفي الممارسة الإكلينيكية. وقد أنشئ أصلاً للاستعمال في التجارب لاختبار فروض معينة تتصل بتأثير الدافع على التعلم، وتحقيقاً لهذا اختبرت عبارات معينة من الاختبار المتعدد الأوجه على أنها تدل على تعبير ظاهر عن القلق، وكانت نتيجة هذا الاختبار (50) عبارة، زيد عليها عبارات إضافية لا تتصل بالقلق ومختارة من مقاييس الضبط الثلاثة (K, F, L)

كذلك نتج عن الجهود التي بذلت لتعديل الاختبار المتعدد الأوجه بحيث يستعمل مع تلاميذ المرحلة الثانوية والطلبة الجدد في الكليات أن أنشئ اختبار منسوجاً للإرشاد النفسي ويتكون هذا الاختبار من (413) عبارة على المفهوم أن يجيب عنها بصواب أو خطأ، ومظم هذه العبارات مأخوذ من اختبار للشخصية المتعددة الأوجه. ويعطي اختبار منسوجاً للإرشاد النفسي درجات في سبعة مجالات هي العلاقات الأسرية، والعلاقات الاجتماعية، والتهبات الانفعالي،

والمسايرة، والتكيف مع الواقع، والمزاج، والقيادة. وعلى الرغم من أن مقاييس لاختبار الإرشاد النفسي تسمى باسم الطرف المقبول، أي الثبات الانفعالي مثلاً بدل من الطرف الآخر وهو الاضطراب الانفعالي، فإنه تتشابه مع مقاييس الاختبار المتعدد الأوجه المسماة بأسمائها المرضية، فمثلاً مقياس (المسايرة) لاختبار الإرشاد النفسية يشبه إلى حد كبير مقياس الانحراف للميكوباتي في الاختبار المتعدد الأوجه، وكذلك مقياس التكيف مع الواقع يشبه كثيراً مقياس الفصام ولهذا الاختبار مقياسان للضبط يشبهان مقاييس للضبط في الاختبار المتعدد الأوجه وتشير بعض الدراسات إلى أن مقاييس هذا الاختبار ذات قدرة إيجابية ولكنها ضعيفة على تمييز للتلاميذ الذين يعرف لهم غير متوافقين أسرياً من غيرهم ممن يتمتعون بتوافق جيد.

ومن الاختبارات التي اعتمدت على الاختبار المتعدد الأوجه اختبار كاليفورنيا النفسي فهو يشق نصف عباراته من الاختبار للمتعدد الأوجه ومن أمثلة مقاييسه: السيطرة، النزعة الاجتماعية، ضبط النفس، الإنجاز عن طريق المسايرة، والإنجاز عن طريق الاستقلال والمرونة، وتعتبر أنستازي هذا الاختبار من أفضل اختبارات للشخص المستعملة حالياً، وقد خضع لدراسات وتحسينات عديدة أعطته هذه المكانة. ولهذا الاختبار صورة للتلاميذ الثانوي هي اختبار كاليفورنيا للشخصية.

عرضنا بشيء من التفصيل اختبار للشخصية المتعددة الأوجه لأنه (النموذج الأمثل) لاختبارات الشخصية التي تعتمد على التقرير الذاتي للمفحوص، إلا أن الاختبارات التي تنتمي إلى هذا النوع كثيرة ومتنوعة إلى

درجة يكاد يستحيل حصرها جميعاً وخاصة في هذا المجال المحدود، لهذا نختار نماذج ممثلة من هذه الاختبارات التي لم تتأقش في الصفحات القليلة الماضية مراعين أن يكون لكل منها ميزته التي تجعله جيداً بهذا العرض الموجز.

قائمة مولي للمشكلات وقد وضعها روس مولي وزملاؤه، ولها عدة صور من مراحل التعليم المتوسطة حتى للجامعة والارشء، ويحدد فيها المفحوص مشكلاته في أحد عشر مجالاً منها الأخلاق والدين، للنواحي المالية والمعيشية، والتوافق للعمل المدرسي، والعلاقات الاجتماعية. وتشير الدرجة المرتفعة إلى الحاجة إلى الإرشاد النفسي، وتتخذ العناصر التي يحددها المفحوص أساساً للمناقشة الفردية والجماعية وقد استخدمت هذه القائمة في عدة دراسات عربية.

اختبار التوافق لـ بل

منذ صدور هذا الاختبار في عام 1934 وهو يستعمل على نطاق واسع. وللصورة الخاصة بالطلاب مصممة للانتقاء السريع لطلاب المدارس الثانوية والكليات بهدف الإرشاد النفسي، وهذه الصورة تعطي درجات توافق في مجالات أربعة هي: الأسرة، للصحة، للناحية الاجتماعية، والناحية العاطفية. وللصورة الأخرى خاصة بالكبار وهي أقل استعمالاً، وقد زيد عليها مجال خامس هو مجال التوافق المهني.

مقياس البورت لدراسة السيطرة والخضوع

يعتبر هذا المقياس من أقوى اختبارات الشخصية المبكرة ويوصف بأنه يقيس للسيطرة والخضوع، لأنه يقيس نزعة للشخص إلى أن يسود ويسيطر على

زملائه، أو أن يسلط ويسيطر عليه في علاقات لمواجهة اليومية. ويبدأ كل عنصر بوصف موجز لموقف من مواقف الحياة اليومية العادية في اجتماع أو في مدرسة أو سيارة عامة، ويذكر أسلوبين أو أربعة لمواجهة الموقف، وعلى المفحوص أن يحدد أي الأساليب يعبر أكثر عن سلوكه. وتختلف الاستجابات في درجة السيطرة والخضوع التي تمثلها وتعطي لها أوزان مناسبة لهذه الدرجة في التصحيح. وفضلاً عن شيوع هذا الاختبار فإنه قد أُر في كثير من الاختبارات التي أنشئت بعده. ومن المهم أن نشير إلى أن السيطرة كانت إحدى السمات التي برزت بوضوح في دراسات التحلي العاملي.

البروفيل الشخصي لـ جورتن

يعتبر هذا الاختبار، وقائمة الشخصية للمؤلف نفسه، من اختبارات الشخصية التي تعتمد على الاختبار الإجباري للتغلب على ما يسمى بالمرغوبة الاجتماعية، أي ميل المفحوص إلى أن يبدو مقبولاً اجتماعياً، وكذلك للإقلال من القموض وبعض الصعوبات الأخرى التي تصاحب عملية تصحيح اختبارات الشخصية.

ويقيس البروفيل أربعة جوانب للشخصية هي: السيطرة، المسئولية، والاتزان الانفعالي، والاجتماعية (أو النزعة الاجتماعية) وهذه الجوانب مستقلة نسبياً، والبروفيل مناسب للاستخدام مع طلاب المدارس الثانوية، والجامعة والراشدين.

ويتكون البروفيل من 18 مجموعة من العبارات الوصفية لشمل كل مجموعة على أربع عبارات وتمثل كل عبارة إحدى سمات الشخصية الأربعة،

المشار إليها من قبل. وتتشابه جملتان من الأربع من حيث أن لها قيمة تفضيلية عالية. أما الجملتان الأخيرتان فتتساويان في القيمة التفضيلية المنخفضة. ويطلب من المفحوص أن يضع علامة على جملة من الجمل الأربعة على أساس أنها تشبهه إلى حد بعيد وعلى جملة أخرى على أساس أنها تشبهه إلى أقل درجة وبهذا لا يستطيع المفحوص أن يستجيب استجابة قبول لجميع العبارات كما يحدث في كثير من مقاييس التقرير الذاتي الأخرى أي أن البروفيل يعتبر أقل قابلية لتشويه من جانب أولئك الذين يحاولون أن يعطوا صورة طيبة عن أنفسهم.

اختبارات التحصيل

اختبارات التحصيل

يقصد بالاختبار التحصيلي الأداة التي تستخدم في قياس المعرفة والفهم والمهارة في مادة دراسية أو تدريبية معينة أو مجموعة من المواد.

وهذه الاختبارات هي أكثر الأدوات شيوعاً في التقويم للتربوي، وهي في بلادنا هي الوسيلة الوحيدة التي تستخدم في توجيه للتلاميذ ولانتقائهم، كما أنها في أغلب الأحوال ما يمكن للمعلم أن يصنعه في ممارسته اليومية للتدريس. ولذلك فإنها تحتل منزلة خاصة بين أدوات التقويم النفسي بجعل من المهم لإيراد هذا الفصل لها بحيث يمكن تناولها بشيء من التفصيل حتى يمكن للمعلم والسلطات التربوية عامة، أن تستفيد من المبادئ التي سوف نتناولها هنا في إعداد هذه الاختبارات.

تحديد الأهداف للتعليمية

أشرنا في الفصل الأول من هذا الكتاب إلى العلاقة بين التقويم النفسي والأهداف التربوية. والواقع أن هذه الأهداف تتفاوت من العبارات الشاملة التي تدل على الأغراض العامة والقيم والفلسفات التربوية إلى العبارات المحددة التي تبين أساليب أداء نوعية أو خاصة. ويشار إلى النوع الثاني في علم النفس التربوي المعاصر بالأهداف التعليمية، وإذا كانت الأهداف التربوية يعوزها اللوضوح في أغلب الأحوال، فإن الأهداف التعليمية أكثر تحديداً لأنها تتناول للنواتج النهائية لعملية التعلم والتعلم في صورة إدارة يمكن ملاحظته، وبالتالي فهي على درجة أكبر من اللوضوح.

ويرى كثير من علماء النفس التربويين المعاصرين أنه من الواجب صياغة الأهداف التعليمية في عبارات صريحة تدل على الأداء النهائي لثلاثة أسباب هي:

1- يؤدي تحديد الأهداف التعليمية تحديداً سلوكياً بهذا الشكل إلى توجيه المعلم في تخطيط عملية للتعليم فيمكنه أن يحدد الخطوات التي يجب أن يسير فيها التلميذ للوصول إلى الهدف، كما يمكن أن ينهض بأعباء جميع الاستجابات التي يجب أن تصدر عن التلميذ للوصول إلى الاستجابات الهامة النهائية.

ومن المؤكد أن أساليب التدريس تعتمد على الأهداف التعليمية والمخلات السلوكية، ومع ذلك فإن من نتائج تخطيط المعلم هذا وبحته قد يكون تعديل الأهداف بتغير مقدار الزمن مثلاً أو المستوى المتوقع للإتقان أو الموضوعات التي يجب تناولها في التدريس، وبالطبع لا يمكن لهذه التغيرات أن تحدث ما لم تتم صياغة الأهداف بصورة صريحة. وفي نفس الوقت لا يستطيع المعلم أن يوائم بين أساليبه في التعليم والفروق في المخلات السلوكية ما لم يحدد بشكل دقيق ما يستطيع التلميذ أدائه قبل للتعليم وما يتوقع منه أدائه بعد التدريس، ومعنى ذلك فإن أهم أسباب وجوب صياغة الأهداف بشكل صريح ما يهيؤه ذلك للمعلم من إرشاد في تخطيط أساليبه في التدريس.

2- يفيد تحديد الأهداف التعليمية في عبارات صريحة في تقويم الأداء ومن المعروف أن الاهتمام بالصياغة الدقيقة للأهداف التعليمية بدأه في الأصل العلماء المهتمون ببناء الاختبارات وتقويم المناهج المدرسية، فقد اكتشفوا أن استخدام العبارات الغامضة في التعبير عن الأهداف يجعل عملية بناء الاختبار ومفرداته عملية صعبة إن لم تكن مستحيلة.

3- تنفيذ ترجمة الأهداف التعليمية في صورة أداء نهائي في توجيه لنتائج التلميذ وجهوده.

وقد أثبت ميجر ومكان أن التلمذ بمن أن يحدث حيث لا يفعل المعلم شيئاً أكثر من إعطاء التلاميذ قائمة بالأهداف التعليمية الواضحة يقوم التلاميذ بالاهتمام بها في تعلمهم لذاتي.

وصف العمل التربوي

يتطلب تحديد الأهداف التعليمية وصفاً دقيقاً للعمل التربوي كما يتطلب تحليلاً تفصيلاً لهذا العمل، وطريقتنا وصف العمل وتحليل العمل نشأنا أول الأمر في ميدان علم النفس الصناعي ثم انتقلنا في السنوات الأخيرة إلى ميدان علم النفس التربوي.

ومن أهم أساليب وصف العمل التربوي طريقة ميجر وطريقة ميلر، وتتطلب طريقة ميجر في وصف العمل، أو بعبارة أخرى تحديد الأهداف التعليمية، خطوات ثلاثة هي:

1- تحديد الأداء النهائي الذي يسعى التعليم أو للتدريس إلى إنتاجه.

2- وصف الظروف والأحوال الهامة التي نتوقع حدوث السلوك فيها.

3- وصف مستوى الجودة الذي يجب أن يصل إليه أداء التلميذ.

فمثلاً قد يكون الهدف التعليمي الذي نسعى إلى تحقيقه هو أن يستطيع التلميذ تحديد العوامل التي أدت إلى حدوث الحرب العالمية الأولى، وهذا هدف مترجم في أداء يمكن ملاحظته ملاحظة صريحة وعلمة في صورة أداء نهائي

يسمى التعليم إلى إنتاجه (الخطوة الأولى)، ثم يمكن أن نعطي للتلميذ قائمة بالعوامل التي تؤدي إلى الأحداث التاريخية الهامة، وتكون هذه القائمة وصفاً لما يسميه ميجر العوامل الهامة التي يحدث في ظلها السلوك النهائي (الخطوة الثانية)، وأخيراً تتحدد المستويات المطلوبة (الخطوة الثالثة). والواقع أن تحديد المستويات هذا ليس من الأمور للقطعة ويتراوح بين الإتيان الكامل وبين أقل مستويات الإتيان. وتشير طريقة ميللر في وصف العمل إلى وجود ثلاث خطوات أيضاً هي:

- 1- ظروف الاستثارة أو ما يسميه indication.
- 2- للتنشيط أو الاستجابة التي يجب أن تصدر عن التلميذ.
- 3- للتغذية الراجعة أو تحديد مستوى ملائمة الاستجابة.

والواقع أن هذه الخطوات لا تختلف كثيراً عن تلك التي أشار إليها ميجر، فظروف الاستثارة هي ما يشير إليه ميجر بالعوامل أو الظروف أو الأحوال الهامة، والتنشيط عند ميللر يقابل ما يسميه ميجر بالأداء النهائي، ويتشابه مفهوم التغذية الراجعة عند الأول مع مفهوم مستوى الأداء عند الثاني.

تحليل العمل التربوي

بعد ما يقوم المعلم بوصف الأعمال التي يتوقع من تلاميذه أداء ما يكون عليه تحليل هذه الأعمال وتصنيفها. والقرص الرئيسي من تحليل العمل هو مساعدة المعلم على تحديد أنسب ظروف التعلم، والتي تختلف من نوع من أنواع التعلم الآخر. فتعلم تكوين المفاهيم أو حل المشكلات أو الابتكار يتطلب ظروفاً تختلف عن اكتساب المهارات الحركية أو حفظ قائمة من الكلمات. وهذا يعني

بالطبع أن مسؤولية المعلم أكبر من مجرد وصف العمل التربوي أو صياغة الأهداف في عبارات سلوكية صريحة، فهي تشمل أيضاً تصنيف هذه الأساليب السلوكية المتضمنة في كل هدف من هذه الأهداف ثم تحديد أنسب شروط التعلم لكل منها.

ويوجد في الوقت الحاضر تصنيفان أساسيان للسلوك لهما أهميتهما التربوية هما تصنيف جاجنيه وتصنيف بلوم. ويشمل تصنيف جاجنيه ثمانية أنواع من التعلم تتدرج من البسيط إلى المركب هي:

- 1- التعلم الإشاري أو ما يسمى في سيكولوجية التعلم بالاشتراط الكلامي وفيه يتم تعلم استجابة شرطية لمثير أو إشارة معينة.
- 2- تعلم المثير، الاستجابة أو ما يسمى بالاشتراط الإجرائي.
- 3- التسلسل أو ما يسمى اكتساب المهارات الحركية (غير اللفظية).
- 4- الارتباط اللغوي.
- 5- التمييز المتعدد.
- 6- تعلم المفاهيم.
- 7- تعلم للمبادئ.
- 8- حل المشكلة أو التفكير.

أما التصنيف الذي وضعه بلوم ومساعدوه للميدان العقلي المعرفي فيشتمل على فئتين كبيرتين هما المعرفة والقدرات والمهارات العقلية بحيث نحصل في النهاية على ستة فئات للأهداف التربوية هي:

1- المعرفة.

- 2- الفهم.
- 3- التطبيق.
- 4- التحليل.
- 5- التركيب.
- 6- التقويم.

وصف المداخلات السلوكية

تتطلب عملية التطعيم وصفاً للمدخلات السلوكية لدى التلاميذ حتى يمكن للمعلم أن يخطط لتنفيذ عملية التدريس، بل وحتى يمكن للسلطات التربوية أن تخطط للعملية التربوية عامة ويتطلب هذا الوصف تحديداً دقيقاً للوضع الراهن لمعارف التلميذ ومعلوماته ومهاراته وطرق تفكيره وقدراته واستعداداته وميوله وديافعه وخصائص شخصيته، قبل أن نسعى إلى إكسابه النماذج السلوكية الجديدة من خلال التعلم والتدريس. وهذه البيانات يجب استخدامها لأغراض الوصف وليس لأغراض التفسير، لأن هذه المفاهيم تصف في الواقع المستوى السلوكي للفرد في لحظة معينة، ولا تستطيع الاختبارات والمقاييس التي نستخدمها في قياس هذه الظواهر أن تفسر لنا أداء الفرد فيها، فلا يمكن مثلاً أن نعرض ضعف الأداء في اختبار للذكاء إلى (ضعف الذكاء) لأن ذلك يصبح نوعاً من الحصول للحاصل ولا يزيد فهمنا للملوك الإنساني.

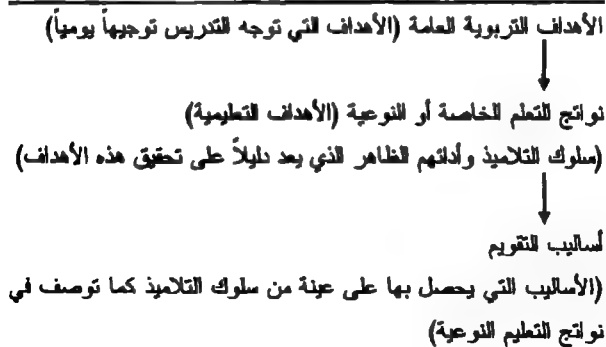
والواقع أن وصف للمدخلات السلوكية ضروري لأي برنامج تربوي. فعملية التطعيم يجب أن تبدأ مما لدى التلاميذ ويجب أن تستمر إلى حيث يستطيعون من أداء نهائي تحت شروط معينة ولمستوى معين. فما أن نحدد أهدافنا يمكننا أن نقرر ما سوف نعلمه وما نفترض أنه لدى التلاميذ بالفعل.

كن للتفاعل بين الأهداف والمداخلات السلوكية أن يؤدي إلى تغيير الأ^ه
ثبت أنها غير واقعية، كما تتفاعل المداخلات السلوكية مع وسائل التدر
اليه، فلحن لا تبدأ تدريسا لتحقيق هدف تربوي معين بمواد تعليمية م
أ وإما قد تستدعي الاستجابات التي سبق تعلمها وخاصة تلك التي تتد
موقف التعلم الجديد.

ثم إن لكل نوع من التعلم مدخلاته السلوكية فتمية سلوك حل الم^ه
لب معرفة بمداخلات سلوكية تختلف عما يتطلبه تنمية إحدى المهام
كية.

علاقة بين الأهداف التعليمية وأساليب التقويم التربوي

يلخص جرونلاند العلاقة بين الأهداف التربوية والأهداف للتعل
الليب للتقويم التربوي على النحو الآتي:



ويتطلب هذا ربط أساليب التقويم مباشرة بنواتج التعليم التي نقومها وبهذا نتيقن من أننا نقوم للتلميذ نحو الأهداف التي أخذناها كإطار لعملية التدريس. وعملية ربط أساليب التقويم بنواتج التعلم النوعية هي في أساسها عملية تحليل منطقي وحكم. وهذه العملية يمكن تيسيرها باستخدام خطة ما من خطط التقويم المنظمة، هذه الخطة يجب أن تشمل على الأكل قائمة بنواتج التعلم المرغوب فيها والأساليب التي تستخدم في تقويم التقدم نحوها. وإليك خطة على سبيل المثال في ميدان علم الأحياء للصف الأول للثانوي:

الأهداف التعليمية ونواتج التعلم النوعية، أساليب التقويم:

1- يعرف التلميذ المصطلحات الشائعة المستخدمة في علم الأحياء حينما:

- يعرف المصطلحات الشائعة في علم الأحياء.
- يميز بين المصطلحات الشائعة على أساس المبني.
- يتعرف على معاني المصطلحات الشائعة حين تستخدم في نص معين.

2- يظهر للتلميذ القدرة على التفكير الناقد حينما:

- تميز بين الحقائق والآراء.
- يشق استنتاجات صحيحة من المعطيات.
- يتعرف على الافتراضات وراء الاستنتاجات.
- يتعرف على لواحي القصور في المعطيات.

3- يؤدي للتلميذ الأعمال الأساسية للتشريح بمهارة حينما:

- يضع العينة في الوضع الصحيح للتشريح.

- يقطع بمهارة دون إفساد العينة التي يدرسها.
- يقسم أجزاء اللبنة في العينة دون إفسادها.
- ينتهي من التشريح في الوقت المحدد.

4- يظهر التلميذ القدرة على تحديد مواضع المعلومات البيولوجية حينما:

- يستخدم كتالوج المكتبة في تحديد المراجع.
- يعرف المصادر المشتركة للمعلومات البيولوجية.
- يستخدم جداول المحتوى والفهرس عند البحث عن للمعلومات في الكتب
- يحدد مدى ملائمة المعلومات لمشكلة محددة.

5- يكون لدى التلميذ اتجاه على نحو الظواهر البيولوجية حينما:

- يؤجل الحكم حتى تتوفر جميع الحقائق أو معظمها.
- يبحث عن علاقات السبب والآخر في البيانات البيولوجية.
- يظهر رغبة في التفسير الجديدة للبيانات البيولوجية.
- يفسر البيانات البيولوجية متحرراً من التحيز.
- يظهر ثقة في البيانات البيولوجية التي يحمل عليها بالأساليب العلمية.

من هذا الجدول نجد ما يلي:

1- نواتج التعلم النوعية كما تصاغ في صورة ملوك للتلاميذ عديدة ومتنوعة بحيث لا توجد أداة تقويم واحدة تعطي تليلاً مباشراً على تحصيل هذه الأهداف التربوية. ولذلك نحتاج إلى جانب الاختبارات للموضوعية إلى قوائم الملاحظة والتقارير القصصية واختبارات المقال.

2- تظهر أهمية الصياغة الواضحة للأهداف ونواتج التعلم من انتقاء أداة التقويم. وفي الواقع حينما تصاغ نواتج التعلم في ضوء سلوك للتلميذ فإنها لا تدل فقط على ماذا نقوم ولكنها توحى أيضاً بكيف نقوم.

مثال:

(يحدد المصطلحات الشائعة) يوضح لنا أسلوب للتقويم الذي يجب أن يستخدمه. فهي تدل على أن التلميذ يجب أن يعطي التعاريف بنفسه وعلى ذلك فإن اختبار المثال من نوع الإجابة القصيرة حيث يعطى للتلاميذ عدة مصطلحات ويطلب منه تعريفها هو الأسلوب الأكثر ملاءمة في تقويم هذا الهدف. أما الاختبار الموضوعي مثلاً من نوع الاختبار من متعدد حيث يطلب من التلاميذ أن يتعرفوا فقط على التعريف لا تتناسب مع تقويم هذا الناتج التطبيقي في صيغته هذه، وإلا فإن الصيغة تصبح (يتعرف التلميذ على معنى المصطلحات الشائعة)، وفي هذه الحالة يمكن استخدام الاختبار الموضوعي. ويتوقف هذا على غرضنا من التقويم، فإذا كنا نريد أن نتأكد من أن التلاميذ يعرفون المصطلحات عن طريق تحديدهم وتعريفهم لها فإن الإجراء الوحيد الملائم هو أن يطلب من التلاميذ أن يعرفوا مصطلحاتهم. والقدرة على التعرف على التعريف الصحيح لا يمكن اعتبارها دليلاً على قدرة التلاميذ على إعطاء التعريف الصحيح ما لم يكن هناك دليل من البحث العلمي يؤكد أن نوعي السلوك مرتبطان ارتباطاً عالياً. وحيث أنه لا يوجد مثل هذا الدليل بعد فإن الإجراء الأحوط هو اختيار أسلوب التقويم الذي يرتبط مباشرة بناتج التعلم النوعي (الهدف التطبيقي).

تحليل محتوى المادة الدراسية

أشرنا إلى أن الاختبار التحصيلي يرتبط بمادة دراسية معينة ثم ندرسها بالفعل، ومعنى ذلك أنه لا يوجد ما يبرر إعداد اختبارات من هذا القبيل تتناول مواد لم يتم تدريسها أو لم تكون محددة كمحتوى لعملية التعليم. ولذلك لا بد في إعداد الاختبار التحصيلي من تمثيل كل ما حظي باهتماما أثناء عملية للتدريس. ومن الطرق الملائمة في تحديد الأصل الإحصائي المناسب الذي تشتق منه عينة السلوك التي تسمى الاختبار التحصيلي أن تعد قائمة بالموضوعات التي تتضمنها هذه المادة بأكبر قدر من التفصيل وفي أبسط صورة ممكنة. ويمكن للباحث في هذا الصدد أن يستفيد بمنهج البحث المعروف باسم (تحليل المحتوى).

تحديد الوزن النسبي للأهداف التعليمية ومحتوى المادة الدراسية

بعد تحليل المحتوى يقوم الباحث بتحديد الوزن النسب لكل موضوع (أو وحدة) من موضوعاته، ومن المحركات المستخدمة في تحديد وزن كل موضوع من موضوعات المحتوى هو الزمن المخصص لتدريس كل منها، إلا أنه في بعض الأحيان لا يكون الوقت المستغرق في تدريس موضوع معين دلالة على أهميته، ولذلك فإن الأوزان النسبية للموضوعات لا بد أن تعكس اهتمام المعلم به. وبالمثل لا بد من المعلم أن يحدد الأوزان النسبية للأهداف التعليمية لمادته. فإذا كان يؤكد في تدريسه وصول التلاميذ إلى التصميمات وتطبيقها مثلاً فوجب أن يعطى هذا الهدف وزناً أكبر. وبعبارة موجزة لا بد أن تعطي لموضوعات المحتوى وللأهداف التعليمية أوزاناً نسبية تعكس الاهتمام الذي تحظى به في عملية التدريس.

وفي تحديد هذه الأوزان قد يعتمد المعلم على خبرته في تدريس المادة وتحليله لمحتواها فوضع قائمة بأهم الموضوعات ثم يصنفها حسب الأهداف التي يرى أن بقيمتها الاختبار. إلا أنه في كثير من الأحيان لا يكتفي الباحث بالاعتماد على خبرته الشخصية وحدها ويلجأ إلى بعض الخبراء لاستشارتهم في تحديد هذه الأوزان النسبية.

إعداد جدول المواصفات

على أساس المعلومات التي يتم تجميعها يتم تحديد مواصفات الاختبار والتي تتضمن الموضوعات التي يجب أن يشملها الاختبار ونواتج التعلم التي يجب اختبارها (أي الأهداف للتعليمية)، والأهمية النسبية (أو الوزن النسبي) للموضوعات والأهداف. وعلى هذا الأساس يتم تحديد عدد الأسئلة التي يتم إعدادها مرتبطة بكل هدف في كل موضوع والطريقة الأكثر سهولة في تحديد مثل هذه المواصفات هي إعداد جدول ثنائي التصنيف يسمى جدول المواصفات توضع فيه الأهداف للتعليمية لفتياً وموضوعات المحتوى رأسياً. والجدول الآتي مثال على ذلك، ومنه يتضح عدد المفردات التي يشملها الاختبار بالنسبة لكل موضوع من موضوعات المحتوى في علاقته بكل هدف من الأهداف للتعليمية.

المحتوى	معرفة المصطلحات الشائعة	معرفة الحقائق النوعية	فهم المبادئ والتوصيات	تطبيق المبادئ والتوصيات	تفسير الصور والرسوم	المجموع الكلي %
الطعام	2	6	2	-	-	10
الملابس	2	6	2	-	-	10
النقل	4	2	2	2	5	15
الاتصال	4	2	2	2	5	15
المسكن	-	-	5	5	-	10
حياة المدينة	4	2	6	8	-	20
حياة القرية	4	2	6	8	-	20
المجموع الكلي %	20	20	25	25	10	100

وقد مثلت خانات هذا الجدول على أساس الوزن النسبي لكل هدف (ف) حورة للنسب المئوية الموجودة في المطر الأخير من الجدول)، والوزن للنسب ال موضوع من موضوعات المحتوى (في صورة النسب المئوية الموجودة في عمود الأخير من الجدول). وكل رقم في كل خانة يدل على النسبة المئوية أسئلة الاختبار التي تخصص لكل ميدان من ميادين المحتوى وفي كل هدف ه أهداف التعليمية. أمثلاً بالنسبة لموضوع الطعام تتطرق 2% من أسئلة الاختبار معرفة المصطلحات الشائعة فيه، 6% لمعرفة الحقائق النوعية، 2% لفه مبادئ والتوصيات، وتتل الخانتان للخليتان أمام موضوع الطعام على ه جود أسئلة فيه حول تطبيق المبادئ والتوصيات وتفسير الصور والرسوم.

والوزن النسبي المعطى لكل ميدان من مبادئ المحتوى، ولكل نوع من الأهداف في جدول المواصفات يدل على النسب المئوية الكلية للأسئلة المخصصة لكل منها. فمثلاً في العمود الأخير نجد أن 10% من الأسئلة تتعلق بموضوع الطعام، 10% بموضوع الملابس، 15% بالنقل، وهكذا. وبالمثل فإن السطر الأخير في الجدول يوضح أن 20% لمعرفة الحقائق النوعية، وهكذا. وبالطبع لا بد من تحديد هذه الأوزان للنسبية قبل إعداد المواصفات كما أشرنا.

تحديد نوع الأسئلة

أشهر أنواع الأسئلة التي تستخدم في ميدان الاختبارات لتحصيلية أسئلة المقال والأسئلة الموضوعية، وأسئلة حل المشكلات (أو الأسئلة للتفسيرية). ونعرض فيما يلي عرضاً موجزاً لأنواع الأسئلة الشائعة في التقويم التربوي الحديث.

أسئلة الاختبار من متعدد

وتعتبر أكثر الأسئلة للموضوعية شيوعاً، وتقيم بكفاءة شديدة النتائج البسيطة للتعليم، بل قد تصبح لقياس بعض النتائج التعليمية المعقدة. ويتكون السؤال من مشكلة (قد تصاغ في صورة سؤال مباشر أو عبارة ناقصة) تسمى الجذر، وقائمة من الحلول المقترحة (قد تشمل كلمات أو أعداداً أو رموزاً أو عبارات) وتسمى البدائل الاختيارية ويطلب من المفحوص قراءة جذر السؤال وقائمة البدائل وانتقاء البديل الصحيح أو الأفضل. والبديل الصحيح في كل سؤال يسمى ببساطة الإجابة بينما تسمى البدائل الاختيارية المحولات. وهي تسمى

كذلك من واقع وظيفتها المقصودة، أي تحويل هؤلاء الذي في شك من الإجابة الصحيحة.

ويستخدم هذا النوع من الأسئلة في قياس كثير من الأهداف التعليمية وخاصة نواتج التعلم في ميدان التفكير واكتساب المعلومات، ومن ذلك معرفة المصطلحات ومعرفة الحقائق للنوعية، ومعرفة المبادئ والقوانين، ومعرفة الطرق والإجراءات، كما تستخدم في قياس الفهم مثل القدرة على تطبيق الحقائق والمبادئ، والقدرة على تفسير علاقات السبب والأثر، والقدرة على تبرير الطرق والإجراءات.

ويتميز هذا النوع بمرونته المشددة، ومحررة من نقائص الأنواع الأخرى من الأسئلة، ومن مشكلة أسلوب الاستجابة. بل إن صياغة البدائل الاختيارية قد تؤدي استخدامها في أغراض التشخيص. ومع ذلك فإن لأسئلة الاختيار من متعدد مشكلاتها وأهمها أنها تقتصر على المستوى اللفظي للأداء (وهي في هذا تشترك مع غيرها من أسئلة الورقة والقلم)، كما أنها لا تصلح لقياس بعض مهارات حل المشكلة في الرياضيات والعلوم، والقدرة على تنظيم الأفكار (وهي في هذا تشترك مع غيرها من أسئلة التعرف على الإجابة). ثم إن لأسئلة الاختيار من متعدد نقصة خاصة بها لا تشترك فيها مع غيرها، وهي صعوبة إيجاد عدد كاف من المحاولات المحبذة وغير الصحيحة في وقت واحد. إلا أن هذه الصعوبة تتناقص مع زيادة الخبرة في بناء مثل هذه الأسئلة.

أسئلة التفسير

يتطلب هذا النوع من الأسئلة ما هو أكثر من الحفظ واكتساب المعلومات الجزئية، أي أنه يستخدم في قياس نواتج التعلم المرتبطة بالعمليات العقلية العليا كالفهم ومهارات التفكير وقدرات حل المشكلة. ويتكون السؤال

التفسيرى من سلسلة من الأسئلة الموضوعية تعتمد على مجموعة مشتركة من البيانات الأولية (المعطيات). وقد تكون هذه المعطيات في صورة مواد مكتوبة أو جداول أو رسوم أو أشكال أو خرائط أو صور. وقد تتخذ الأسئلة المرتبطة صوراً مختلفة ولكنها في الغالب تأخذ صورة الاختيار من متعدد.

أسئلة المزوجة

تتكون أسئلة المزوجة التقليدية من عمودين متوازيين يحتوي كل منهما على مجموعة من العبارات أو الرموز أو الكلمات. وتسمى المفردات في العمود الذي نلشد له المزوجة بالمقدمات، وتسمى المفردات التي نختار منها الاستجابات response.

وعادة ما تكون أسس مزوجة الاستجابات المقدمات واضحة في ذاتها، وفي بعض الأحيان تتطلب توضيحاً في التعليمات. ويستخدم هذا النوع في قياس الحقائق التي تستند إلى الدداعي البسيط مثل الشخصيات والإنجازات، للتواريخ والأحداث التاريخية، المصطلحات والتعاريف، القواعد والأمثلة، الرموز والمفاهيم، المؤلفون ومؤلفاتهم، الكلمات الأجنبية ومقابلاتها بالعربية، الآلات واستخداماتها، النبات وتصنيفاته، الحيوان وتصنيفه، الأشياء وأسمائها، الأعضاء ووظائفها. وتتميز بأنها تقيس بدقة وسرعة مقدراً كبيراً من الحقائق المرتبطة في زمن قصير نسبياً، بالإضالة إلى أنها سهلة الصنع. ومع ذلك فإن لهذا النوع مشاكله وأهمها أنه يقتصر على قياس الحقائق التي تعتمد على التذكر والدداعي البسيط. وكذلك يصعب على الباحث إيجاد مادة متجانسة لها أهميتها في المحتوى الدراسي يمكن أن تصاغ منها قائمة المقدمات.

أسئلة الإجابة القصيرة (الاستدعاء)

يتطلب هذا النوع من الأسئلة الذي يعرف في بعض المؤلفات باسم أسئلة التكملة - أن ينتج المفحوص استجابته - وليس مجرد التعرف عليها كما هو الحال في الأنواع السابقة. ويمكن القول أن هذا النوع قد يتطلب إجابة قصيرة إذا عرضت المشكلة في صورة سؤال مباشر، أو تكملة إذا عرضت في صورة عبارة ناقصة. وتوجد أنواع أخرى من هذا النوع منها: أسئلة إعداد القوائم والتي تسمى أحياناً اختبارات المقال ذات التعليمات المجددة و أسئلة التمثيل أو التناسب، وأسئلة المشكلات، وأسئلة للتحعين (كأن يطلب من التلميذ تعيين أجزاء رسم نبات أو حيوان). ويستخدم هذا النوع في قياس للنواتج البسيطة نسبياً للتعلم مثل معرفة المصطلحات، ومعرفة الحقائق النوعية، ومعرفة المبادئ، ومعرفة الطرق والإجراءات، والتفسيرات البسيطة، والقدرة على حل المسائل في الرياضيات والعلوم. وهذا النوع ربما يسبب بساطة نتائج التعلم التي تقوسها (الاستدعاء في أغلب الأحوال)، بالإضافة إلى تحرره من أثر التخمين السائد في اختبارات التعرف بوجه عام. ومع ذلك فقد تكون أكبر مشكلاته شيوعاً صعوبة تصحيحه.

أسئلة الصواب والخطأ (أو أسئلة التبدلين)

وتتطلب اختيار إجابة واحدة من إجابتين كالحكم على العبارة بالصواب أو الخطأ، أو الإجابة على السؤال بنعم أو لا، أو الحكم على العبارة بأنها تدل على رأي أو حقيقة، أو تقدير عبارة بالموافقة أو المعارضة. ويستخدم هذا النوع في قياس نتائج التعلم التمييزي البسيط ومع ذلك فيمكن أن تقيد في قياس الفهم (كالتعرف على علاقات السبب والآخر). وقد تكون أكثر مشكلات هذا النوع من الأسئلة خطورة تأثيرها البالغ بالتخمين.

أسئلة الترتيب

وفيها يقوم الملموح بإعادة ترتيب خطوات أو أحداث أو إجراءات أو تواريخ في تسلسل طبيعي أو منطقي.

أسئلة المقال

رغم الاستخدام الواسع للأسئلة الموضوعية التي أشرنا إليها لا زالت توجد نتائج هامة للتعليم لا يصلح لها إلا سؤال، المقال، ومنها القدرة على عرض وتنظيم وتكامل الأفكار، والقدرة على التعبير الكتابي، والقدرة على إعطاء التفسيرات والتطبيقات للمعلومات لا مجرد استدعاؤها أو التعرف عليها.

ومن أهم خصائص أسئلة المقال حرية الاستجابة، وفي هذا جدواها كمقياس للتفصيل المعقد، وفيه أيضاً تكمن صعوبات التصحيح التي تجعل منها أداة أقل كفاءة وكفاءة في قياس الحقائق والمعلومات. ومعنى ذلك أن أسئلة المقال يجب أن تستخدم حيث لا تصلح الاختبارات للموضوعية.

الاختبارات العملية

تستخدم هذه الاختبارات كوسائل موضوعية لتقدير الكفاءة التي يؤدي بها العمل (كالكتابة على الآلة الكاتبة) أو قيادة السيارات. وتنقسم هذه الاختبارات إلى ثلاث أنواع.

أ- لاختبارات التعرف: وتتطلب من المفحوص التعرف على الخصائص الأساسية للأداء (كأن تعزف مقطوعة موسيقية ويطلب من المفحوص تحديد الأخطاء أو للفرق في الأداء أو للتفسير)، أو تحديد الأجزاء

التي تتألف منها إحدى الآلات: أو اختبار الآلة أو الجهاز المناسب لعمل معين، أو تحديد العينات أو تصنيف الأشياء، أو انتقاء عمل أدبي أو فني ممتاز.

ب- الاختبارات التي تتضمن مواقف تشبه المواقف الطبيعية: فهي تهدف إلى قياس الأنشطة الأساسية في العمل، وتسمى أحياناً لاختبارات للنماذج المصغرة.

ج- اختبار عينة العمل: وهو عبارة عن محاولة (مضبوطة) أو (مقلنة) في الظروف الواقعية للعمل، وتنقسم هذه الاختبارات إلى نوعين أساسيين: أولهما الاختبارات التي يسهل فيها التمييز بين الصواب والخطأ في الأداء وبالتالي يمكن تصحيحها بسهولة مثل التصويب، والأداء العضلي في التربية البدنية، والتجميع الميكانيكي، والكتابة على الآلة الكاتبة. وللنوع الثاني هو الاختبارات التي تعتمد على حكم المراقبين والمفحصين لتقويم الأداء وإعطاء درجة أو رتبة، ومن أمثلة ذلك محاولة سامي إبراهيم على (تقنين) لاختبارات المزف على البيانو، ويتطلب هذا النوع استخدام مقاييس التقدير أو قوائم للملاحظة.

الاختبارات الشفوية

الاختبار الشفوي هو مزيج من اختبار المقال والاختبار العملي، وله فائدته في دراسة العمليات العقلية التي يستخدمها المفحوص في الإجابة على أسئلة معينة، ولذلك فهو أداة لافعة في تشخيص الصعوبات. بل إنه في بعض الأحوال هو الأسلوب الأوحى في تقويم المفحوص كما هو الحال في قياس صغار التلاميذ (في الصفوف الأولى من المرحلة الابتدائية)، وفي قياس بعض نتائج التعلم للغوي (كالقراءة الجهرية).

تحديد عدد الأسئلة

يتحدد عدد الأسئلة التي يتكون منها الاختبار بمقدار الزمن الميسر له. وكثير من الاختبارات التحصيلية يتحدد زمنه بحوالي (50) دقيقة (زمن الحصة المدرسية العادية) وبعض جداول الامتحانات تحدد فترات أطول تصل إلى ثلاث ساعات. وعلى وجه العموم كلما طال الزمن وزاد عدد الأسئلة كان الاختبار أكثر ثباتاً. ومن النادر من الوجهة العملية أن نعد اختباراً يستغرق أكثر من ثلاث ساعات.

ومن المفيد كما أشرنا أن نعتبر مجموعة الأسئلة التي يكون منها الاختبار عينة من أصل الفتراضي لكل الأسئلة الممكنة التي يمكن أن تستخدم في الاختبار. فمدرس الصف الخامس الابتدائي قد يحصل على (100) كلمة لاختبار المفردات اللغوية وذلك بانتقاء الكلمة الخامسة من قائمة تحوي على (500) كلمة درمت أثناء العام الدراسي. وفي هذه الحالة تصبح الـ (500) كلمة هي الأصل الذي تؤخذ منه عينة الاختبار، وهو أصل حقيقي ومحدد المعالم. إلا أنه في معظم الاختبارات لا نجد مثل هذا التحديد الحاسم للأصل الإحصائي العام للأسئلة، فمثلاً لا يوجد حد أقصى لعدد المشكلات والمسائل المختلفة التي يمكن إعدادها في اختبارات الجبر.

وبالطبع كلما اتسع مدى الموضوع أو القدرات التي يهدف الاختبار إلى قياسها زاد عدد الأصل الإحصائي. وحجم هذا الأصل يمثل الحد الأقصى لحجم عينة الاختبار. ومعنى ذلك أن أسئلة الاختبار لا يمكن أن تكون أكثر عدداً من الأصل التي تشتق منه، إلا أن حجم هذا الأصل لا يصنع حداً أدنى لحجم العينة. فالأصل الذي يتألف من (1000) سؤال يمكن أن تكون العينة المشتقة منه (10) أسئلة أو (50) سؤالاً أو (200) سؤال ولكما زاد حجم الأصل قل تجانسه، أي أصبح يحتوي ميادين مختلفة وشبه مستقلة من المعرفة والقدرة. وللوصول إلى

نتائج دقيقة يجب أن تكون عينة الاختبار المشتقة من أصل غير متجانس كبير نسبياً من تلك التي تشتق من أصل متجانس. ومع ذلك فلا توجد قاعدة عامة تحدد العلاقة من حجم الأصل والعينة.

ويوجد في الوقت الحاضر اتجاه يؤكد استخدام عدد قليل وكاف من الأسئلة بحيث يعطى معظم المفحوصين وقتاً يحاولون فيه جميع الأسئلة بمعدلاتهم الخاصة العادية. واشتد هذا الاتجاه ثلاثة أسباب:

- 1- أن سرعة الاستجابة ليست هدفاً أولياً للتعلم في معظم المقررات الدراسية وبالتالي لا تعد مؤشراً صادقاً للتصنيف لأنه في كثير من ميادين التصنيف لا ترتبط السرعة والدقة ارتباطاً عالياً. وبالطبع فقد تكون اختبارات السرعة أكثر ملاءمة في بعض الحالات، إلا أن هذه المواقف هي الاستثناء وليس القاعدة.
- 2- أن قلق الاختبار قد يكون على درجة كبيرة من الحدة - حتى في الاختبارات غير الموقوتة - حين يتزايد التركيز على العمل بسرعة.
- 3- أن الاستخدام الكفء للاختبار التحصيلي - الذي يصنعه المعلم - يتطلب أن معظم التلاميذ يستجيبون له كله حتى يستخدم نتائجهم في أغراض التشخيص.

تعليم التحصيل في القراءة في المدرسة الابتدائية

قد تكون القدرة على القراءة أهم القدرات التي تؤكد لها المدرسة الابتدائية. وفي الصفوف الأولى من المرحلة الابتدائية يحل تعليم القراءة كمهارة أساسية الجزء الأكبر من العمل المدرسي، ومع انتقال التلميذ إلى الصفوف العليا تتفرع دراسته إلى مجالات متنوعة ويقل الوقت المخصص لتعليم القراءة، ومع ذلك فإن لها أهميتها طوال المرحلة الابتدائية. وبعد ما يتقن التلميذ بعض المهارات الأساسية في القراءة، قد تتعدّل تعديلات طفيفة أهداف تدريس

القراءة بحيث تؤكد للتغيرات الناتجة في المهارات التي نقوسها. ومن ناحية أخرى فإنه لا زال من الأهداف الأساسية للتعليم فهم معاني الكلمات وفهم الجمل والفقرات الطويلة، وهي جميعاً لها أهميتها في جميع المستويات، ومعنى هذا أن أي اقتراحات تتعلق بتقويم أهداف للقراءة في أحد صفوف المرحلة ينطبق على كل المرحلة. وفيما يلي مثال لأهداف نوعية لتعليم للقراءة في الصف الأول الابتدائي:

الفرض للعلم	الهدف المرحلي (في نصف العام)	الأهداف النوعية
أن يستطيع للفرد المتعلم قراءة للغة القومية قراءة فعالة	أ- أن يستطيع التلميذ قراءة القسم الأول من كتاب القراءة بسرعة وفهم.	1- أن يستطيع الطفل أن ينطق بكلمات مثل مدرسة، الحديقة، ورد، حينما يراها مكتوبة.
		2- أن يستطيع الطفل أن يربط كلمات مثل ولد، طائرة، بيت وبين الصور التي تعبر عنها.
		3- أن يستطيع الطفل أن يجيب على أسئلة تدور حول مواد جديدة يقرأها من مختارات من كتاب للقراءة.
ب- أن يستطيع الطفل القراءة الجهرية بسهولة مع بعض التعبير.	4- أن يستطيع الطفل أن يغير من نبران الصوت للتمييز بين الجملة الخبرية والجملة الاستهامية.	

ومن الواضح أنه لو كان للوجب هو تقويم هذه الأهداف النوعية جميعاً فيجب أن تستخدم أساليب مختلفة. فالهدفان (1، 4) يمكن تقويمهما ببعض صور الأداء الشفوي. بينما الهدفان (2، 3) يمكن تقويمهما بهذه الطريقة أو باستخدام بعض المواد المكتوبة. وبسبب الفوائد الكثيرة الخاصة باستخدام الاختبارات الموضوعية في المدرسة الابتدائية فإن على المعلم أن يختار لقياس تحصيل الهدف النوعي رقم (2) طريقة المزوجة بين صور الأشياء والكلمات التي تدل عليها وفي هذا المستوى يجب أن يعطى لجميع التلاميذ تعليمات شفوية تقرر أنهم يجب عليهم أن يرسموا خطأ يصل بين الكلمة والصورة التي تدل على معنى الكلمة. ومثل هذا السؤال يعطى قياساً مباشراً لقدرة الطفل على ربط كل كلمة بالصورة الملائمة. والملائكة البسيطة بين الهدف وهذا الاختبار تدل على فائدة صياغة الأهداف في عبارات تدل على سلوك ممكن ملاحظته.

وبالنسبة لتقويم الهدف (3) فإن المعلم قد يطلب من تلاميذه الانتقال إلى مواد جديدة وفراستها قراءة صامتة. وبعد هذه القراءة توجه إليهم أسئلة يطلب منهم كتابة إجاباتها عليها وتتعلق هذه الأسئلة بمحتوى للفقرة. وهذا الاختبار قد يأخذ صورة أن يطلب من التلاميذ أن يكتبوا جملة في الإجابة على السؤال، وهو صورة مبسطة لاختبار المقال، إلا أن التلاميذ في هذا المستوى لا نتوقع لهم أن تكون عندهم سهولة كبيرة في الكتابة، ولذلك فإن بعض الاختبارات الموضوعية ينتج قياساً أكثر صدقاً لفهم القراءة عندهم مثل أسئلة الاختبار من متعدد مثل:

يوم العيد ذهبنا إلى: الحقل

للمدرسة

للحديقة

ومرة أخرى يجب أن تكون تعليمات هذه الأسئلة شفوية وتطلب من التلميذ أن يضع خطأً أو يرسم دائرة حول الإجابة الصحيحة للسؤال.

وحيث أن القراءة هي مهارة اتصال، فإن جميع أهداف للتعليم في هذا الموضوع تهتم أساساً بالفهم، وخاصة فهم معاني الكلمات. فمن المتوقع من التلميذ أن يشرح معنى كلمة، أو يختار مرادفاً لها، أو يربط بينها وبين الصورة أو الشيء الذي تصفه. كما تربط الأهداف الأخرى بينهم الجمل أو الفقرات.

ومع تقدم الطفل في صفوف المرحلة الابتدائية فإن للتأكيد على التحصيل في القراءة يتزايد اهتمامه بالملامح المختلفة وجوانب مختلفة من الفهم، فتتطلب اختبارات الفهم من التلميذ أكثر من مجرد استدعاء أجزاء من معلومات الحقائق والتفاصيل.

تقويم التحصيل في الحساب

تتزايد أهمية المهارات الحسابية والفهم الحسابي في الحضارة المعاصرة تزايداً مستمراً. ومن المهم للتلاميذ المرحلة الابتدائية أن يتقنوا تطبيقاً جيداً في المهارات الأساسية في الحساب وفي عناصر التفكير الكمي. وللمساعدة في تحقيق هذا الهدف من المهم للمعلم أن يقوم بتقويم مستمر لتقديم كل تلميذ، وهذا مهم على وجه الخصوص في الحساب حيث كل مهارة لاحقة تعتمد على المهارة السابقة، وحيث يفضل للتلميذ إذا فشل في إتقان أي من المعلومات الأساسية مثل:

للغرض العام	الهدف المرحلي (في نصف العام)	الأهداف للتوعية
أن يستطيع الشخص المستعلم حل مشكلات العدد والحساب	أ- أن يستطيع التلميذ يد حتى للرقم (100)	1- أن يستطيع التلميذ أن يعد ويسجل العدد الصحيح لأشياء أو حروف توضيحها للصور في حدود (10) وحدات.
	ب- أن يعرف التلميذ حقائق الجمع الأساسية	2- أن يستطيع الطفل العد إلى الرقم (100) بزيادة واحد.
	ج- أن يستطيع التلميذ فهم المصطلحات الحسابية الخاصة	3- أن يستطيع التلميذ الإجابة على مسائل الجمع لأزواج من الأعداد التي تتكون من رقم واحد بحيث يكون مجموعها أقل من (10).
		4- أن يستطيع التلميذ شرح معنى مصطلحات مثل جمع وطرح ومجموع وأكثر وأقل.
		5- أن يستطيع التلميذ أن يختار اختصاراً صحيحاً الأشياء التي تتحدد بمصطلحات مثل الأول والثاني والثالث... إلخ

وجميع هذه الأهداف للتوعية الخمسة يمكن تقويمها إما شفويًا أو تحريريًا، فالهدف الأول يمكن قياسه باستخدام أسئلة بسيطة يعطى فيها للطفل أعداداً مختلفة من الأشياء عليه أن يعدها.

تعليمات (شفوية): أكتب على الخط الموضوع أسفل كل صورة العدد الذي يدل على عدد الحروف في كل منها:

د	د	ص	ص	ح	ح	5	5	5
د	ص	ص	ص	ح	ح	5	5	
	ص	ص		ح	ح	5	5	5

ويمكن تقويم الهدف الثاني بسؤال التلميذ أن يعد جهرياً أو أن يكتب الأعداد من 1 إلى 100.

والهدف الثالث يمكن قياسه بكفاءة أكثر بمجرد سؤال التلميذ أن يكتب إجاباته على عدد من المسائل التي تتضمن الجمع البسيط:

$$\begin{array}{ccccccccc} 6 & 1 & 5 & 7 & 3 & & & & \\ & & & & & 3+ & 8+ & 4+ & 2+ & 4+ \end{array}$$

والهدف الرابع يمكن تقويم بسؤال التلميذ مباشرة أن يشرح هذه المصطلحات شفوية أو تحريرية. ويوجد تعديل طفيف لهذا الهدف بحيث يتطلب من التلميذ أن يتعرف على صور أشياء أو عمليات يمكن أن توصف لفظياً كأن يعطى له رسم لعدد من الكرات متقلوبة الحجم ويطلب منه وضع خط تحت الكرة الأصغر.

والهدف الخامس يمكن تقويمه باستخدام أسئلة مماثلة، كأن يعطى له رسم لعدد من الكرات المتساوية ويطلب منه وضع خط تحت الكرة الرابعة مثلاً.

إن أنواع الأسئلة التي ذكرناها حتى الآن تتعلق أساساً بقياس معلومات التلميذ وتذكره لحقائق خاصة أو مصطلحات نوعية. إلا أن المعلم قد يهتم بأنماط أخرى من القدرات وخاصة تلك التي ذكرناها في تصنيف الأهداف التربوية (البloom) تحت عناوين مثل الفهم والتطبيق.

مثال لترجمة الأعداد والمفاهيم إحداها إلى الآخر.

1- ما الذي يعطي نفس الإجابة التي يعطيها 4×5 ؟

أ- $4 + 4 + 4 + 4 + 4$

ب- $5 + 5 + 5$

ت- $4 + 5$

ث- $4 \div 5$

2- $\frac{3}{4}$ تساوي

أ- $\frac{6}{8}$ ب- $\frac{5}{8}$ ج- $\frac{4}{8}$ د- $\frac{3}{8}$

ومن أهم الأهداف التعليمية التي نحتاج إلى تقويمها في اختبارات الحساب هي القدرة على تطبيق معرفة للمبادئ والأساليب أو الإجراءات (وتقاس باختبارات المسائل الحسابية من مختلف الأنواع).

تقويم التحصيل في العلوم

مع زيادة التأكيد على العلم في منهج المدرسة الابتدائية ككل، نجد من الطبيعي أن تسمى المدرسة الابتدائية إلى زيادة مدى ما يتضمنه برنامجها من العلوم. والأهداف التي تسعى تحقيقها يجب أن تكون أكثر من مجرد معرفة حقائق معينة أو مصطلحات أو ح فظ مبادئ وقوانين. وإنما يجب أن تتضمن قدرات مثل القدرة على إجراء تجارب بسيطة، واشتقاق المبادئ والتعميمات من البيانات أو إعطاء أمثلة للمبادئ المختلفة وتطبيق مثل هذه المبادئ في شرح النتائج أو التنبؤ بها كما يجب أن تتعلق الأهداف باتجاهات وميول معينة.

وحيث أن أغلب مطمي المدرسة الابتدائية يعوزهم الإعداد السليم في العلوم فإنهم يحتاجون إلى المساعدة في تحديد أنماط التعلم التي تنتج عن مقرر العلوم تحديداً واضحاً وفي الحصول على أمثلة خاصة من القدرات المختلفة ومن الأساليب المفيدة في هذا الصدد إعداد ولقاء واختبارات جيدة وأسئلة جيدة في الاختبارات، فكثير من القدرات يمكن للتدليل عليها بطريقة جيدة عن طريق تطبيقها في مواقف تعد أسئلة اختبار أمثلة لها أو مواقف تتحول إلى أسئلة اختبار. ومعرفة الطريقة التي يمكن بها تقوم هذه القدرات يساعد في توضيح المطى الحقيقي للقدرة عند العلم وتشجعه على تدريسها.

وفيما يلي مثال:

الأهداف النوعية	الهدف المرحلي	الفرض العام
1- أن يستطيع التلميذ أن يذكر أسباب ضرورة المحافظة على مصادرها من الماء والأرض والنبات والحيوان.	أ- أن يفهم التلميذ أهمية المصادر الطبيعية مثل الماء والأرض والنبات والحيوان.	لأولاً: على المواطن المتعلم أن يقدر مصدرة الأمة.
2- أن يستطيع التلميذ تحديد الطرق المختلفة للمحافظة على المصادر الطبيعية.	1- أن يستطيع التلميذ أن يتعرف على إسهامات علماء مثل باستير وديسون ومولك ومركوني.	ثانياً: يقدر المواطن المتعلم التقدم العلمي بإسهامه في الرفاهية العامة.
1- أن يستطيع التلميذ أن يحل مشكلات العلم بالوصول إلى النتائج من الأدلة والشواهد.	أ- أن يفهم التلميذ الطريقة العلمية في حل المشكلات.	ثالثاً: تتوفر لدى المواطن المتعلم وسائل دفاع ضد الادعاء
2- عندما يطلب من التلميذ أن يتخذ قرارات في مواقف فرضية في (مجالات العلوم وغيرها) فإنه يصل إلى تلك التي تدعمها الشواهد والأدلة.	ب- أن يصل التلميذ إلى قرارات على أسس الأدلة والشواهد سواء في مسائل العلم أو المسائل غير المتعلقة بالعلوم	

إن الهدفين النوعيين المشتقين من الغرض العام الأول يهتمان بمعرفة للمبادئ والطرق، ويمكن تقويم الهدف الأول بمجرد سؤال التلميذ أن يذكر قائمة إما شفويّاً أو كتابةً، بعدد معين من الأسباب التي تتعلق بالمحافظة على كل مصدر من المصادر الطبيعية المذكورة.

ويمكن للهدف اللّثاني أن يصاغ على نحو يشبه الهدف الأول، وبالتالي يمكن تقويمه بأن نطلب من التلميذ أن يذكر قائمة شفوية ويمكن أيضاً تقويم هذا الهدف باستخدام الاختبارات الموضوعية.

مثال: حتى يستطيع الفلاح أن يحصل من أرضه على محصول وفير عليه أن يزرع في حقله:

- أ- نفس المحصول عاماً بعد عام.
- ب- محصولاً واحداً لثلاث سنوات متوالية ثم محصولاً آخر السنوات الثلاثة التالية.
- ج- محصولاً واحداً لخمس سنوات ثم يليه محصول آخر للسنوات الخمس التالية.

ويمكن صياغة الهدف رقم (1) بحيث يسمح للتلاميذ بانتقاء الأسباب الهامة للمحافظة على المصادر الطبيعية، وعلى هذا النحو يمكن تقويمه باستخدام أسئلة الاختبار من متعدد. أما الهدف النوعي المعتمد على الغرض للعام الثاني

فيمكن لاختباره باستخدام أسئلة بسيطة مثل أسئلة المزاجية بين أنواع الإسهامات التي قام بها مختلف العلماء.

والأهداف المشتقة من الغرض العام للثالث أصعب في تقييمها نوعاً ما. ويمكن أن نفيذ في ذلك الاختبارات الشفوية واختبارات المقابل، حيث يعرض على التلميذ مواقف مشكلة معينة ثم يطلب منه إعطاء الاستنتاجات والطول المناسبة. كما يمكن تقييم هذه الأهداف باستخدام أسئلة الاختبار من متعدد من النوع التفسيري كما هو مبين في المثال الآتي:

مثال: قام أربعة تلاميذ بالاشتراك في تجربة لاستنبات الفول كما يلي:

- قام المعتز بمسح بعض حبوب الفول بقطعة من القطن ثم تركها جافة.
- قامت مها بمسح بعض حبوب الفول بقطعة من القطن ثم تركتها مبللة.
- قام خالد بوضع بعض حبوب الفول في طين جاف وتركها جافة.
- قام أحمد بوضع بعض حبوب الفول في طين مبلل وتركه مبللاً.

ومن نتائج هذه التجربة حلول تلاميذ للفصل الحكم على أي النتائج الأكثرية حول الصحيح:

- أ- تثبت حبوب الفول إذا تركت فوق قطعة من القطن الجاف.
- ب- تثبت الحبوب إذا تركت فوق قطعة من القطن المبلل.
- ج- تثبت الحبوب إذا تركت في الطين الجاف.
- د- تثبت الحبوب إذا تركت في الطين المبلل.

هـ- تحتاج الحبوب إلى أن تظل مبللة حتى تثبت.

و- تحتاج الحبوب إلى أن تظل جافة حتى تثبت.

ز- تحتاج الحبوب إلى التربة والماء حتى تثبت.

المراجع

- 1- التقويم النفسي، د. فؤاد أبو حطب، د. سيد أحمد عثمان، مكتبة الأنجلو المصرية - القاهرة.
- 2- الاختبارات الامقاطية، سيد محمد غنيم، هدى عبد الحميد، النهضة العربية.
- 3- التشخيص النفسي، ج1، سيد محمد غنيم، هدى عبد الحميد، النهضة العربية.
- 4- التشخيص النفسي، ج2، سيد محمد غنيم، هدى عبد الحميد، النهضة العربية.

0518323

مجمع البحار
مكتبة مجمع البحرين للمطبوعات

962-9-366521
 962-9-366521
 E-mail: Mo_pub@hotmail.com

دار اقتصادي للنشر والتوزيع
الطبعة العربية السعودية - الرياض
الطبعة: 009662176134 فاكس: 009662176134
Email: Dep-techsupport@edconl

دار صفاء للطباعة والنشر الإلكتروني
عنوان: شارع السليمان - مجمع المصطفى (بغداد)
هاتف: 00964 78221921 - فاكس: 00964 78221921
www.darsafa.com E-mail: darsafa@darsafa.com